RUNX1/AML1 WORKS IN TISSUE STEM CELLS AT THE CROSSROADS OF SEVERAL EPITHELIAL CANCERS

A Dissertation
Presented to the Faculty of the Graduate School
of Cornell University
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy

by
Cornelia Johanna Franziska Scheitz

January 2013

RUNX1/AML1 WORKS IN TISSUE STEM CELLS AT THE CROSSROADS OF SEVERAL EPITHELIAL CANCERS

Cornelia Johanna Franziska Scheitz, Ph.D.
Cornell University 2013

The transcription factor Runx1 is essential for hematopoietic stem cell emergence and hair follicle stem cell activation. Further it has a distinct role in leukemia development and is strongly expressed in mouse squamous cell carcinoma. Here I investigated Runx1 expression in other epithelial stem cells and its role in malignancies as a means to understand its potential as a global stem cell and cancer regulator. Utilizing lineage tracing I showed that Runx1 is a stem cell intrinsic gene in oral tissue. Additionally, I discovered that Runx1 is expressed in intestinal tissue stem and progenitor cells in a pattern reminiscent of Runx1 expression in hair follicles.

Using an inducible epithelial knockout system, lineage tracing, and tissue culture manipulation, I analyzed the timing and mechanisms by which Runx1 influences cancer formation specifically in the skin and oral tissue. I discovered that Runx1 loss in the skin prevents tumor initiation, impairs tumor promotion, and causes regression of fully developed tumors. Moreover, by tracing the fate of Runx1 expressing hair follicle stem cells I demonstrated that they are at the origin of squamous cell carcinoma. Besides skin tumors I showed that mouse oral tumors develop in a Runx1-dependent manner.

To elucidate the mechanism behind these phenotypes I use cell lines derived from human and mouse tissues treated with shRNA against Runx1. Skin, oral and ovarian cell lines of either mouse or human origin fail to grow in the absence of Runx1. By adding activated Stat3 I can rescue this phenotype *in vitro*. Utilizing primary keratinocytes I uncovered that Runx1 directly binds and represses SOCS3 and SOCS4 resulting in continuous activation of Stat3, while not affecting its expression.

Collectively, my *in vitro* results, staining of clinical human samples, and meta-analysis of published microarray data suggest that the mechanism of Runx1 in cancer development is conserved in several human cancers. Overall, my results support a role for Runx1 as a global stem cell regulator and provide mechanistic insight into its function in solid tumor formation.

BIOGRAPHICAL SKETCH

Cornelia Scheitz was born on May 16th 1985 in Berlin, Germany. For 7 years she attended a Bilingual Gymnasium with focus on English education but including Russian, French, Latin, and Afrikaans. Early on she was interested in genetics and basic life sciences. To follow this interest she switched to a science focused Gymnasium and in 2004 she completed her Abitur among the top 1% of her class with focus on Biology, Mathematics, and English.

From 2004 to 2007 Cornelia attended Jacobs University Bremen in Germany for a Bachelor of Science in biochemistry and cell biology. During those years she also devoted her time to build the community of the young university by starting traditions and being active in various clubs such as debating and ballroom dancing. Cornelia managed several events on campus like international Model United Nations conferences and the Graduation Ball in 2007. In the summers of 2005 and 2006 she completed two industry internships. Her second internship at the LADR GmbH investigating polymorphisms in patients with metal allergies lead to an invited talk at the 2nd WOREAL Conference in 2007.

The same year in June Cornelia graduated from Jacobs University and moved on to complete as Masters degree in Human Molecular Genetics at Imperial College London. In the six months of research she worked with Dr. Mario Falchi on detecting CNV data on monozygotic twin SNP data. In September 2008 Cornelia received her MSc degree with Distinction as the best of her class. In London she

also worked half-time in Quality Assurance and Language testing for the international game development company Square Enix Ltd.

In September 2008 Cornelia joined Dr. Andrew Clark's lab at Cornell University where she started to work on variance in lipid profiles of geographically distinct populations of *Drosophila Melanogaster*. In 2009 she started her PhD degree in the field of Genetics and Development and joined the lab of Dr. Tudorita Tumbar investigating the effect of a hair follicle stem cell factor on skin cancer. During her time at Cornell Cornelia taught "Introduction to Genetics" at the Auburn maximum-security correctional facility for 2 semesters as part of the Cornell Prison Education Program.

This is dedicated to my loving mother. For your love and support and for letting me find my own path.

And to my best friend and partner: Jonas. For always being there for me, motivating me, and allowing me to keep an open mind.

ACKNOWLEDGMENTS

I would like to first thank my chairs Doina Tumbar and Andy Clark for their trust, support, critical feedback and fruitful discussions over the years. Both of them taught me how to think critically and analytically. Each of them allowed me to acquire vastly different skill sets and introduced me to different approaches to solve scientific problems. Doina showed me the tricks and tweaks of how to write an appealing paper, a skill that I will have to develop further. Lastly, I want to thank Andy for giving me the opportunity to come to Cornell University and join the PhD program. I would also like to thank my committee members Bob Weiss and Eric Alani for their advice and guidance beyond committee meetings.

Additionally, I want to thank all of my wonderful colleagues, past and present, in the Clark and Tumbar Labs. Especially, Shu and Song for helping me to start my endeavor in mouse skin leading to the research presented in this thesis. Tae Lee for his enthusiasm and commitment to our project and all the late-shifts we shared.

And, of course, I would also like to thank my friends near and far. Especially Nadja, Laura and Nora.

Lastly, I would like to thank my entire family for their ongoing support and faith in me. Xenia, Jonas, Mimi, Annette, Petra, Renate; without you this would not have been possible.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

4-OHT	4-Hydroxytamoxifen
BCC	Basal cell carcinoma
Bu	Bulge
ChIP	Chromatin immunoprecipitation
CSC	Cancer Stem cell
DAG	Diacylglycerol
DAPC	Discriminant analysis of principle components
DMBA	9,10-Dimethyl- 1,2-benzanthracene
E	Embryonic day
FA	Fatty acid
HF	Hair follicle
HFSC	Hair follicle stem cell
IFE	Interfollicular epidermis
iKO	inducible KO
INF	Infundibulum
IRS	Inner root sheath
K5	Keratin 5
K8	Keratin 8
K10	Keratin 10
K14	Keratin 14
K15	Keratin 15
KD	Knockdown
KI	Knockin
KO	Knockout
lysoPC	lysophillic phosphatidylethanolamine
lysoPE	lysophillic phosphatidylcholine

M	Matrix
ORS	Outer root sheath
OSM	Oncostatin M
PA	Phosphatidic acid
PC	Phosphatidylcholine
PCC	Pearson Correlation Coefficient
PCo	Principle component
PD	Postnatal day
PE	Phosphatidylethanolamine
PG	Phosphatidylglycerol
PI	Phosphatidylinositol
PS	Phosphatidylserine
MMC	Modulated modularity clustering
REML	Restricted Maximum Likelihood
S	Saturated
SC	Stem Cell
SCC	Squamous cell carcinoma
SG	Sebaceous gland
SLE	Systemic lupus erythematosus
SNP	Single nucleotide polymorphism
TAG	Triacylglycerol
TPA	12-O-tetradecanoylphorbol-13-acetate
TM	Tamoxifen
US	Unsaturated
WT	Wild-type
X-gal	5-bromo-4-chloro-3-indoxyl-beta-D- galactopyranoside

PREFACE

My thesis is a collection of two separate projects that I conducted during my time at Cornell. Since this was not planned as such, this preface is aimed at explaining the background to this development.

I joined Cornell as a non-degree graduate student in the laboratory of Dr. Clark working on fly lipid metabolism. I got the opportunity to join the Genetics and Development PhD program and was very intrigued by a project presented by Dr. Tumbar investigating the role of Runx1 in skin cancer development. In August 2009 I ultimately decided to do a co-advised PhD based on this project, but adding a sequencing component to it, that would combine both advisors' expertise. Unfortunately we could not obtain sufficient funding to implement this experimental part. Thus, after completing my study on Runx1's role in cancer development I decided to return full-time to Dr. Clarks lab and continue my work on lipid profiles in *D. melanogaster*.

The title and first two Chapters of this thesis report my work with Dr. Tumbar. Chapter 3 is dedicated to my work in Dr. Clark's laboratory.

CHAPTER 1[1]

NEW INSIGHTS INTO RUNX1 CONTROL IN EPITHELIA STEM CELL BIOLOGY AND PATHOLOGY

1.1 Overview

The transcription factor Runx1 has been studied in leukemia and blood for decades, but recently it has been also implicated in epithelial biology and pathology. Particularly in mouse skin Runx1 modulates Wnt signaling levels thereby regulating timely induction of hair follicle specification, proper maturation of the emerging adult hair follicle stem cells in embryogenesis, and timely stem cell (SC) activation during adult homeostasis. Moreover, Runx1 acts as a tumor promoter in mouse skin squamous tumor formation and maintenance, likely by repressing p21 and promoting Stat3 activation. Similarly, Runx1 is essential for oral epithelium tumorigenesis mediated in mice by Ras, and for growth of three kinds of human epithelial cancer cells. In contrast, Runx1 has a tumor suppressor function in the mouse intestine and shows tumor subtype specific behavior in human breast cancer. Multiple studies revealed Runx1 SNPs to be associated with human cancers and autoimmune disease. With this information as background, the field is poised for functional and mechanistic

[1] This work is in press "Scheitz CJS and Tumbar T (2012) New Insights into Runx1 Control in Epithelia Stem Cell Biology and Pathology. J Cell Biochem" and is reprinted here with permission. The author contributions are as follows: Cornelia Scheitz conducted the literature research, created the panels and wrote the manuscript. Dr. Tudorita Tumbar assisted in manuscript preparation.

studies to elucidate the role of Runx1 in formation and/or progression of epithelial-based human disease.

1.2 The transcription factor Runx1/AML1

More than two decades ago the gene *Runx 1* (*AML1*) was commonly found at chromosome translocation points in acute myeloid leukemia (AML) (Miyoshi et al., 1991). Since then its requirement for adult hematopoietic stem cell emergence and fate specification, and its potential role in leukemia formation drove sustained efforts to characterize its structure and function. Moreover Runx1 plays a role in neural progenitor cell proliferation. Details on Runx1 function in the hematopoietic system, leukemia, and nervous system development have been discussed elsewhere and are beyond the scope of this review (Chen et al., 2009; Mangan and Speck,2011; Kanaykina et al., 2010). This review focuses on recent findings pertaining to Runx1 function and molecular interactions in epithelial tissue homeostasis and disease biology, particularly skin, oral and intestinal tissues.

All Runx family proteins (Runx1, Runx2, Runx3) contain the characteristic DNA binding *runt* domain that was first identified in *Drosophila* (reviewed by Friedman, 2009). Runx1 forms a complex with the core binding factor β (CBFβ) resulting in a stable *runt* domain conformation and thus DNA binding (Yan et al., 2004). Additionally, *Runx1* has transcriptional activation and inhibitory domains allowing it to bind to a plethora of co-factors, such as the co-repressor mSin3A

and the potent co-activator p300, modulating its gene expression regulatory effects (reviewed by Friedman, 2009).

Runx1 has two promoters driving the expression of 3 isoforms (Ghozi et al., 1996): the distal promoter P1 and the proximal promoter P2. P1 controls the longest isoform Runx1c. Runx1a is the shortest and Runx1b the most expressed form, and the proximal P2 promoter drives expression of both of them. All Runx1 isoforms are expressed in the hematopoietic system at different stages of development (Challen and Goodell, 2010), but little is known about their expression in other tissues. Notably, Runx1a lacks the protein binding domains but binds DNA strongly and thus potentially acts as an inhibitor to Runx1b and c. In particular it has been implicated in affecting differentiation of neural crest SCs (Kanaykina et al., 2010). To date there is no data on Runx1a in the skin, but in keratinocytes Runx1b, but not Runx1c, is detected (Ortt et al., 2008). Accordingly, Runx1b is strongly expressed in FACS isolated hair follicle stem cells (HFSCs), but not other basal layer skin cells (Osorio et al., 2008). For the remainder of this review we will consider Runx1 in general, without respect to the isoforms.

In this review we first summarize Runx1's role in mouse and human HF development and homeostasis, followed by its implication as a stem cell gene in the skin and oral tissue, and possibly in the intestine. Second we focus on Runx1's implications in epithelial disease, considering autoimmune disorders and various epithelial cancers, ranging from skin to endometrial cancers.

1.3 Runx1 in hair follicle stem cells

1.3.1 Hair cycle development and homeostasis

The hair follicle (HF) is the best-studied system for Runx1's role in epithelial tissue SCs. The HF morphogenesis occurs in pre- and post-natal life and is comprised of a permanent and a temporary portion. The permanent portion contains the bulge (Bu), which harbors the stem cells, the sebaceous gland (SG) that produces the oil necessary to lubricate the hair canal, and the infundibulum (Inf) through which the hair shafts exits the skin epidermal surface. The bulge stem cells are part of the outer root sheath (ORS) and are marked by surface expression of CD34 (Morris et al., 2004).

In the mouse, hair formation starts at around embryonic day (E) 14.5 and proceeds through the developmental stages of placode, germ, peg, and bulbous peg (Figure 1-1). Then the full HF develops by growing downward into the dermis and generating the inner, terminally differentiated layers of HFs: inner root sheath (IRS) and hair shaft (Schneider et al., 2009). Starting at about postnatal day (PD) 17, morphogenesis is followed by three-week long hair cycles spaced by more and more prolonged periods of quiescence (Figure 1-1 and Figure 2-1). During each cycle the HF undergoes three stages of remodeling. Catagen, the regression phase, characterized by apoptosis of the bulb is followed by the quiescent phase or telogen (Schneider et al., 2009). This is followed by anagen, the growth phase. At the telogen to anagen transition some HFSCs leave the bulge to differentiate and proliferate and form the matrix (M), a population of

short-lived transit amplifying progenitor cells that give rise to all the differentiated, inner layers of the hair follicles throughout anagen. During early anagen the bulge cells divide on average 3 times without further migration and replenish the HFSC pool (Waghmare et al., 2008;Zhang et al., 2009). The synchrony of hair follicle cycle phases and hair morphogenetic stages, and the relative abundance of HFSCs that can be isolated from the skin, makes hair follicle an ideal system to study tissue stem cell biology *in vivo* (Tumbar et al., 2004). The hair follicle served as a platform to begin elucidating the role of Runx1 in epithelial stem cells and cancer.

1.3.2 Runx1 plays a distinct role in hair follicle epithelium and in skin dermis

Runx1 expression in the HF was first identified by Levanon et al. (2001) using E15.5 whisker follicles of reporter mice containing LacZ inserted in frame in the Runx1 genomic locus (Figure 1-1). Closer inspection of pelage HFs at E12.5 revealed the first, rare Runx1 expressing cells in the developing epidermis (Osorio et al., 2011). At placode formation (E14.5) rare epithelial X-gal labeled cells are located in the center of emerging placode structures, and remain present in the upper region during the hair germ stage. As HF structures advance to pegs, more labeled cells are observed mostly in the center of follicles, probably corresponding to the preliminary differentiated HF lineages (cortex, IRS) (Raveh et al., 2006; Osorio et al., 2011). At later stage pegs, Runx1 is also strongly expressed in the infundibulum protruding into the epidermis (Raveh et al., 2006; Osorio et al., 2008, 2011). Remarkably, the Runx1 positive epithelial cells in hair

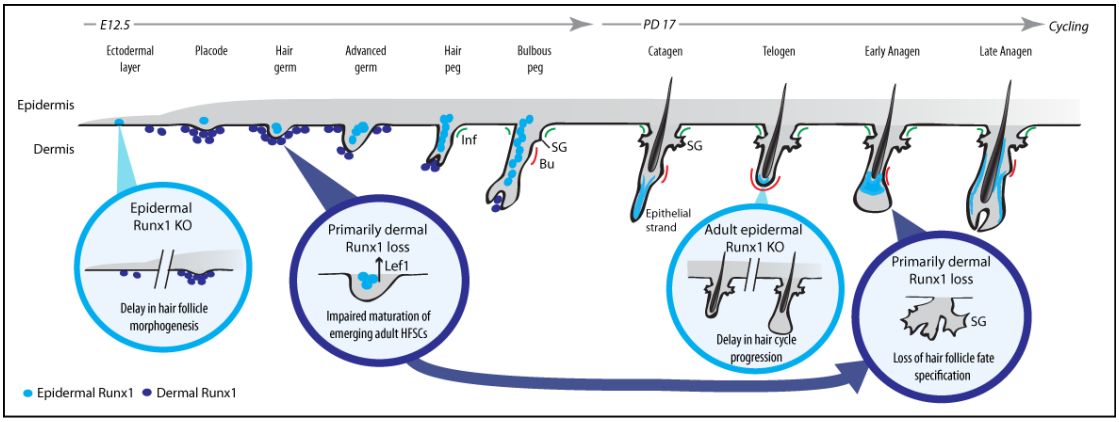


Figure 1-1: Runx1 expression pattern

Summary of the Runx1 expression pattern during hair follicle morphogenesis and the hair cycle. Bubbles summarize the effect of Runx1 loss in specific layers and at specific stages.

placodes contribute to all layers of HF during morphogenesis and adult homeostasis, and thus these cells are pre-cursors of adults HFSCs (Osorio et al., 2011). On the other hand, most epithelial Runx1 expressing cells of later developmental stages (germ, bulbous peg) are short-lived and contribute less often to adult HFs (Osorio et al., 2011).

In telogen Runx1 is expressed weakly in the lower bulge and the hair germ, the cells that proliferate first upon telogen-anagen transition (Raveh et al., 2006; Osorio et al., 2008). Runx1 expression is extended to include the top of the bulge, and the ORS in early anagen. During full anagen the inner differentiated cortex and cuticle layers express Runx1 in addition to the lower bulge cells and the ORS. In catagen Runx1 is not observed in the bulge but in the ORS below it, reaching its peak during apotosis and remaining relatively high in the secondary germ (Raveh et al., 2006; Osorio et al., 2008). Similar to HF precursor cells, the adult bulge cells expressing Runx1 are HFSCs since they contribute long-term to HF homeostasis as well as to oncogenic and physical injury repair (Scheitz et al., 2012). In an initial study, induced epithelial knockout of Runx1 was only found to affect hair structure causing the mouse pelage to look ruffled, a phenotype that can be attributed to Runx1 expression in the cortex (Raveh et al., 2006). Under closer inspection, however, Runx1 loss during morphogenesis and during adulthood causes a temporary delay in morphogenesis and hair cycle at the telogen to anagen transition (Osorio et al., 2008, 2011). This is eventually overcome through injuries or with age (Osorio et al., 2008; Hoi et al., 2010), but it cannot be

ruled out that spontaneous hair cycle progression is caused by micro injuries obtained from normal activities such as grooming. Nonetheless it is clear that both HF establishment and cycling rely on SC migration and proliferation. Given that Runx1 is expressed in SCs during both stages, a failure to initiate migration and/or proliferation in Runx1 knockout skin could explain the observed phenotypes. This hypothesis is supported by the fact that Runx1 expression precedes proliferation in bulge and germ cells of adults HFs (Osorio et al., 2008) and *in vitro* Runx1 is required for keratinocytes proliferation and migration (Osorio et al., 2008, 2011; Hoi et al., 2010).

The fact that injury overcomes the roadblock caused by Runx1 loss implies that different pathways are required for injury response compared to HF homeostasis and HFSC proliferation. While the homeostasis trigger is specifically localized to the HFSCs, injury causes a broad proliferation response in the IFE and the HFs including surprisingly a robust Runx1 expression (Scheitz et al., 2012).

Remarkably, mesenchymal Runx1 loss during HF morphogenesis has a much more profound effect on HF integrity than epithelial loss has. Runx1-expressing cells emerge in the dermal layer directly underlying the epidermis starting at E14.5. This dermal population increases by E17.5 and forms clusters under the hair placode, hair germ, and in the dermal condensates of emerging HFs. By birth this dermal expression diminishes drastically and by the beginning of the adult hair cycle is completely lost (Raveh et al., 2006; Osorio et al., 2008). Using $Runx1^{CreER/fl}$ mice, Osorio et al. (2011) knocked out Runx1 at E12.5-14.5. Runx1-

CreER activation showed low efficiency in the epidermis but strong efficiency in the dermis. Since these mice are not viable, Osorio et al. (2011) showed through skin grafting that in the embryonic mesenchyme Runx1 regulates cell fate specification in the adult HF even after normal protein expression has ceased. The hair follicles form normally in morphogenesis but in the first hair cycle, when adult stem cells generate the differentiated hair lineages, they are converted to enormous sebaceous cysts. Notably, in wild-type mice Runx1 expression in the dermis has long been lost at this point. This short, strong and specific expression of Runx1 in the dermis during embryogenesis has a long-lasting, irreversible, and profound effect on adult hair follicle structure, in what appears as discussed later to be maturation of the emerging adult HFSCs.

Overall, Runx1's phenotypic reach in the epidermis is narrow and temporally confined within close distance of its expression. On the other hand, dermal Runx1 is far-reaching, outlasting and drastically expanding beyond its expression. But while we have made good progress in understanding the role of Runx1in HF development and homeostasis, we only begin to understand the mechanisms involved downstream and Runx1's place in the network of HF factors.

1.3.3 Runx1 is part of a network of factors regulating hair follicle fates

Protein interactions and signaling in the HF are complex (Lee and Tumbar, 2012) and thus the list of potential partners for Runx1 is long. To date Runx1 appears to interact with Lef1 and the Wnt signaling pathway, as well as p63 and Stat3.

The first indication that Lef1 and Runx1 might interact in hair cycle regulation was revealed using a micro-array study of wax depilated hair cycle induction. mRNAs of both Runx1 and Lef1 are up-regulated in late anagen and binding sites for both are found in genes highly expressed in the hair cortex (Ishimatsu-Tsuji et al., 2005). Co-incidentally, epidermal Runx1 knockout skin displayed reduced Lef1 protein levels in both epidermis and dermis and reduced Wnt signaling. Specifically, mRNAs of Wnt activators were up-regulated while mRNAs of Wnt inhibitors were down-regulated by the Runx1 loss. While it is not clear if Runx1 works directly on the promoters of these genes, Runx1's role as a context-dependent transcriptional activator or repressor (Friedman, 2009) could explain this effect. These Wnt regulatory genes included secreted molecules that could explain the paracrine effect of Runx1 loss in the epidermis. Since Lef1 is a not only a mediator but also a direct target of the Wnt pathway (Lee and Tumbar, 2012), it is not surprising that its levels are regulated by Runx1 action on the Wnt pathway. Strikingly, Wnt signaling and Lef1 levels in both HF layers and dermis that are not expressing Runx1 (Osorio et al., 2011) were up-regulated when Runx1 loss was targeted mainly to the dermis via the Runx1-CreER. The opposite effect of Runx1 loss in the epidermis versus dermis on Lef1 and Wnt signaling underscores the context-dependent action of Runx1. Lef1 elevated levels persists into adulthood, a stage when Runx1 expression in the mesenchyme has already been lost in wild-type animals (Osorio et al., 2011). This causes the misspecification of HFSCs fate to SG, as discussed in the previous section and described by Merrill et al. (2001). p63 is so far the only identified upstream

regulator of Runx1 in the HF. Ortt et al. (2008) convincingly show that intronic ΔNp63, but not TAp63, binding to *Runx1* promoter positively regulates Runx1 expression, specifically the *Runx1b* isoform. Most likely both proteins are already co-expressed at E14.5 in the developing epithelium. At this stage Ortt et al. (2008) show broad ΔNp63 expression but fail to see Runx1 in the epithelium; later on Osorio et al. (2011) showed Runx1 to be in fact present in the HF rudiments. In adult mouse HFs Runx1 and p63 partially overlap during anagen in the lower ORS and the differentiating matrix cells. To clarify the functional significance of this regulation, site-directed mutagenesis at the binding site in keratinocytes could yield further insights. Additionally, experiments are required to analyze if the direct regulation of Runx1 by ΔNp63 is conserved across tissues and/or species because previous screens in transformed human cells lines did not uncover this relationship, which could indicate that this mechanism is specific to mice or to skin.

Lastly, Stat3 loss has a remarkably similar phenotype as Runx1 loss (Sano et al., 1999; 2000). While the HFs develops normally, hair cycling and wound healing are impaired in Stat3 KO mice (Sano et al., 1999). In Runx1 KO HFs Stat3 is up-regulated, possibly compensating for Runx1 loss and thus suggesting that Runx1 is found up-stream of Stat3 (Osorio et al., 2008). Indeed Scheitz et al. (2012) show in keratinocytes and skin cancer cells that Runx1 is essential for maintaining active Stat3 by repressing cytokine-signaling suppressors (SOCS). Strikingly, both Runx1 and Stat3 affect normal hair cycling (Osorio et al., 2008;

Sano et al., 2000), however, externally induced hair cycling appears to function independently from both factors. This strongly suggests that the mechanism for Runx1-dependent Stat3-regulation discovered *in vitro* translates to *in vivo* HF biology.

It is likely that in the future we will be able to identify more partners of Runx1 in this network that is regulating HF biology.

1.3.4 Runx1 in human hair

All the findings acquired in mice begged the question of the role of Runx1 in normal and diseased human skin. Importantly, when Runx1 expression was compared in mice and men similar patterns emerged in the hair follicle at different hair cycle stages (Osorio et al., 2008; Raveh et al., 2006; Soma et al., 2006; Hoi et al., 2010). Human Runx1 expression partially overlaps with that of human keratin associated protein 5 (hKAP5.1). In accordance with the presence of putative Runx1 binding regions in the promoter of hKAP5.1, human cultured keratinocytes show increased hKAP5.1 levels upon Runx1 overexpression (Soma et al., 2006). However, more functional studies in human keratinocytes or *ex vivo* cultures are needed to verify protein interactions and molecular functions of candidate factors for human HF homeostasis. In the case of Runx1 it appears that at least its expression pattern is conserved from mouse to humans, giving us hope that this will also be true for its functional consequences.

1.4 Runx1 in other epithelial tissue stem cells

Runx1 expression has been observed in other stratified epithelial organs similar to skin such as nails and keratinized pads (Raveh et al., 2006), but its potential role or overlap with putative SCs other than hair follicle and blood has been unclear. Recent evidence suggests that Runx1 may be present in other epithelial SCs. Scheitz et al. (2012) show that Runx1 expressing cells are present in the basal layer of the oral epithelium. Lineage tracing of these cells in Runx1CreER; tdTomato mice shows that some cells remain in the basal layer for at least 18.5 weeks and give rise to the differentiated epithelium (Scheitz et al., 2012), hence fulfilling the stem cell definition of long-term self-renewal and differentiation. Moreover, Runx1 was detected in some LGR5+ cells of the intestinal SC region and also showed high levels in the differentiated epithelial cells of the crypt, a pattern that resembles that observed in the HF (Osorio et al., 2008; Scheitz et al., 2012). Runx1 promoter activity is also found at the base of colon crypts (Scheitz et al., 2012). Fijnemann et al. (2012) present some evidence that in the absence of Runx1 expression of SC specific genes increases, supporting a putative regulatory role for Runx1 in the colon SCs. However, conclusive lineage tracing data to demonstrate the lineage potential of Runx1 expressing cells in the intestine and colon awaits generation of a Runx1-CreER knockin mouse line that can efficiently be activated for recombination in these tissue (Scheitz et al 2012). On the whole, we are just beginning to discover and understand Runx1 biology in tissue function and maintenance.

1.5 Runx1 in autoimmune diseases

Due to its widespread activation upon injury Runx1's role in the skin may well extend beyond HF growth regulation (Osorio et al., 2008; Scheitz et al., 2012). Indeed SNP-association studies have identified Runx1 as a potential candidate in several autoimmune diseases, which commonly show skin symptoms (summarized in Table 1-1 at the end of this Chapter). It is interesting to note that these associations originate from SNPs within Runx1 binding sites, which may prevent Runx1 from regulating expression of its target gene. Systemic lupus erythematosus (SLE) was the first autoimmune disease with a suggested Runx1 function. A study by Prokunina et al. (2002) identified an intronic SNP in PDCD1 (programmed cell death 1) associated with SLE that disrupts a Runx1 binding site within an enhancer element. Although the directionality of the effect could not be established due to large patient variation, the authors showed that with this intronic SNP PDCD1 fails to bind Runx1 resulting in altered PDCD1 expression (Prokunina et al., 2002). No further functional studies were done, but it is known that PDCD1 regulates T-cell function and regulates self-antigen tolerance (Fife and Pauken, 2011) and thus the axis of Runx1-PDCD1 is a good candidate for autoimmune disease. Rheumatoid arthritis is the only autoimmune disease where functionally independent SNPs within *Runx1* and its binding sites have been identified (Tokuhiro et al., 2003). A SNP within a Runx1 binding site in SLC22A4, a membrane-transporter, is associated with rheumatoid arthritis and causes increased suppression of SLC22A4 due to increased Runx1 binding

(Tokuhiro et al., 2003). Expression of SLC22A4 is limited to the hematological and immunological tissues but the functional consequences of its mis-expression are not known.

While both SLE and rheumatoid arthritis have distinct skin phenotypes, patients are affected primarily by other problems, such as joint pain. Psoriasis, on the other hand, is a skin specific inflammatory disease featuring increased dermal vasculature, T-cell invasion, and overall immune-mediated keratinocyte hyper proliferation (Griffi and Barker, 2007). A family based study by Helms et al. (2003) identified a causative SNP close to the sodium/hydrogen exchanger SLC9A3R1 that disrupts a Runx1 binding site causing loss of Runx1 binding and ceased *SLC9A3R1* transcriptional activity. Given their observation that *SLC9A3R1* is expressed strongly in inactive but not active T-cells, the authors speculate that loss of Runx1 binding causes T-cell activation through a lack of SLC9A3R1. The fact that disease-associated SNPs have not been identified in Runx1 itself indicates that Runx1 must remain functional, while its target SLC9A3R1 does not. Overall, it seems plausible that SLC9A3R1 is regulated by Runx1 because in addition to the SNP-associated site this gene locus contains 32 predicted Runx1 binding sites in humans, 18 of which are conserved to Rhesus macaque and 3 to Mouse.

Functional studies have not been performed yet for many of these SNP associations, but they serve as a solid basis for future investigation of Runx1 role in human disease. To take our knowledge to the next level we need to derive

functional connections with biological significance. In this respect, we know most about Runx1 role in the field of cancer biology. Below we summarize SNP associations and cancer-related functional biological studies in human and in model organisms.

1.6 Runx1 in solid tumors

1.6.1 Biological and mechanistic evidence in skin cancer

Because of its context-dependent function in cell proliferation, apoptosis, and cell differentiation, Runx1 may act in different epithelial cells and tissues as either a tumor suppressor or as an oncogene. Whereas a tumor suppressor function has been suggested recently in intestinal cancer (Fijneman et al., 2012), the first evidence of an oncogenic function is derived from fibroblasts where Runx1 overexpression causes oncogenic transformation (Kurokawa et al., 1996). This is in line with the observed block of proliferation in keratinocytes upon Runx1 loss (Osorio et al., 2008) and to date there is significant evidence that Runx1 is required for tumor formation in the skin. Human skin squamous cell carcinoma (SCC) over-expresses Runx1 and loss of this protein prevents SCC formation in mouse skin (Hoi et al., 2010). In fact, Runx1 is required for tumor initiation in the HFSCs and based on lineage tracing the majority of papilloma (benign squamous tumors which are precursors of SCC in mice) originate from Runx1 expressing HFSCs (Scheitz et al., 2012). Runx1 is crucial for proliferation in normal keratinocytes, while during tumor growth loss of Runx1 can be temporarily overcome with the help of other proliferative agents, such as 12-O-

tetradecanoylphorbol-13-acetate (TPA) (Scheitz et al., 2012). More importantly, Runx1 loss in fully developed mouse tumors leads to significant tumor regression and similarly primary human skin SCC cell lines cannot survive without Runx1 (Scheitz et al., 2012). Hence, Runx1 is required for tumor and cancer cell growth and survival and is a prime potential target for SCC prevention and treatment. Although these data clearly show the dependence of tumors on Runx1, an oncogenic function in the skin has not been clearly demonstrated. In fact, ectopic expression of Runx1 in non-stem cells during injury is insufficient to render those cells tumorigenic, attesting to the importance of the epigenetic context to the Runx1 role in cancer.

Scheitz et al. (2012) also extended the skin phenotype to head and neck SCC, which are suppressed in mice in the absence of Runx1. Moreover, oral human SCCs cell lines require Runx1 for their growth. In the future, Runx1 contribution to esophageal, cervical and lung SCC development and maintenance should be explored, since Runx1 is expressed there at high levels according to Oncomine data (Scheitz et al., 2012). However, even at the current state, the remedy of a Runx1 inhibitor in skin SCC treatment cannot be ignored. SCC is the second most common cancer affecting 0.8 million new patients annually in the US alone (Alam and Ratner, 2001). Current treatment is often disfiguring since SCCs occur frequently on the face and require lengthy and expensive surgery. A non-invasive treatment that could be applied by the patient himself would not only benefit the patient but also reduce health care costs. Mechanistically, Runx1 regulates several genes that have been shown to influence cancer formation in various

tissues, namely p21 and Stat3 (Figure 1-2). p21 is significantly up-regulated in the HFSCs upon Runx1 loss (Osorio et al., 2008) and double knockout of Runx1 and p21 in mouse keratinocytes rescues cell proliferation *in vitro* (Hoi et al., 2010). Hence, Runx1 may mediate its tumor promoting phenotype partially through repressing p21, already expressed in the cells of tumor origin. Additionally, Runx1 prevents expression of the suppressors of cytokine signaling 3 and 4 (SOCS3 and 4) by binding to their promoters, allowing for activation of Stat3 by phosphorylation through the Jak/Stat pathway (Scheitz et al., 2012). The combination of Stat3 activation and p21 repression makes Runx1 a central player in SCC formation. Both genes are involved in other cancers and it remains to be seen if Runx1 is up-stream in those cases as well. Additionally, this connection brings up the question if Runx1 is a broader cancer regulator than previously envisioned. In the following section we will review the evidence that has been collected in other epithelial tissues to give cues for future cancer research.

1.6.2 The role of Runx1 in other solid tumors

Similar to the findings in autoimmune diseases Runx1 has been associated with various cancers in SNP studies (summarized in Table 1-1 at the end of this Chapter). In prostate cancer SNPs within *Runx1* are associated with an increased risk of cancer progression and lymph node metastasis leading to a worse patient prognosis (Huang et al., 2011). Additionally, Runx1 expression increases with pathological stage (Yeh et al., 2009) implying that Runx1 has an oncogenic function and that the SNPs boost or misdirect its activity. In particular, SNPs in

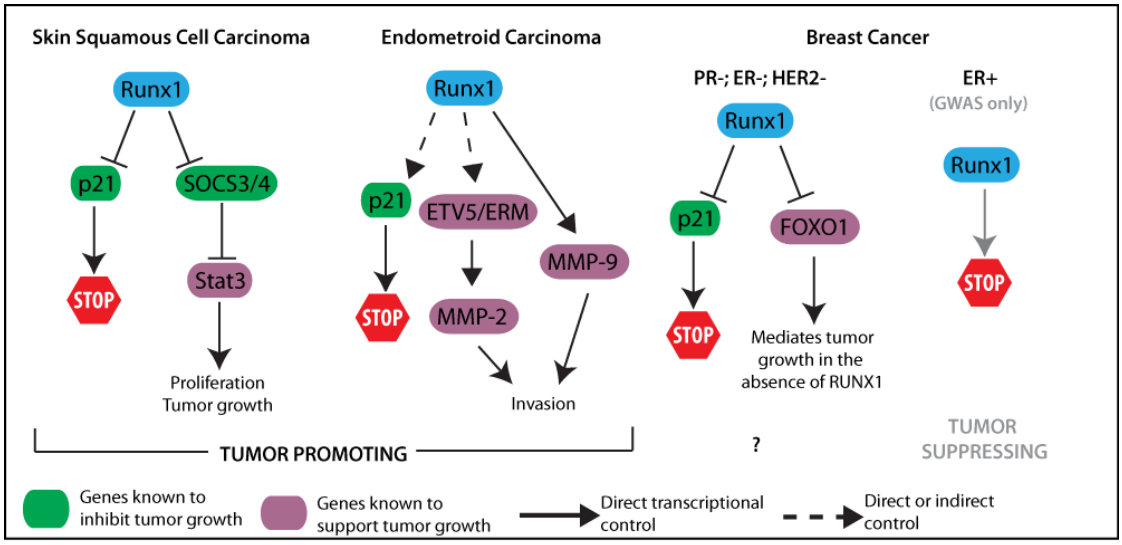


Figure 1-2: Known pathways of Runx1's function in three cancer types

Runx1 also correlated with SNPs in Stat3; Yeh et al. (2009) suggest that in this context both genes regulate expression of EZH2. Although these findings were not tested yet *in vivo*, they provide some intriguing additional evidence that Runx1 can affect tumor development through Stat3.

Colon and rectal cancer were significantly associated with tagging SNPs for Runx1 in a study of two Caucasian case-control cohorts (Slattery et al., 2011). Similar to previous studies of these cancers, single SNPs do not withstand stringent statistical testing due to small effect sizes. However, Slattery et al. (2011) show that simultaneous occurrence of SNPs in *eIF43*, *Runx1* and *Runx3* represents a significant combined risk for colon and rectal cancer. Similarly, combinations of *Runx1* SNPs were associated with CIMP+, MSI+, Kras2 and TP53 colon cancer subtypes and its interactions with SNPs in Smad3, Smad7, BMPR1B, BMPR1A, and TGFβR1 strongly increases the risk for colon and rectal cancer. This study does not conduct biological experiments to identify the functional consequences of these SNPs, yet it suggests that Runx1's phenotype is the result of complex network interactions. At the same time a computational study of cancer microarrays has identified Runx1 positively associated with colon cancer development (Pradhan et al., 2012). Strikingly, a study in mice gives reason to assume that Runx1 can also act as a suppressor of intestinal cancers. Both wild-type and Apc^{Min} mice develop significantly more colon and small intestine tumors in the absence of Runx1 (Fijneman et al., 2012). Notably, both skin SCC and intestinal cancers have been shown to originate from the tissue SCs (Barker et al., 2009; Lapouge et al., 2011; White et al., 2011; Scheitz et al., 2012). Thus Runx1

loss is affecting both SC populations. Still it seems that at least in mice its effect on tumor formation lies on opposite ends of the spectrum. In humans Runx1 seems to be strongly over-expressed in SCCs and colon cancers supporting an oncogenic function, while in intestinal cancers Oncomine analysis uncovered Runx1 as neither up- nor down-regulated (Scheitz et al., 2012). At this point it is not clear if the tumor-suppressor function is conserved across species or intestinal cancer subtypes.

Microarray and immunofluorescence data on primary human tumors, along with loss of Runx1 function studies on cultured cancer cells suggest that at least head and neck and skin SCCs have a conserved Runx1 function, but in breast cancer there is more than one mechanism at play. While in triple negative breast cancer (ER-;PR-;HER2-) Runx1 was either present or compensated by FOXO1 up-regulation (Wang et al., 2011), in oestrogen-receptor-positive (ER+) breast carcinomas Runx1 (or CBFβ) is lost in 4 out of 37 cases (Banerji et al., 2012) (Figure 1-2). Breast epithelial cells (MCF10A) have a molecular expression profile similar to triple negative cancers. Within this *in vitro* model genes have multiple verified Runx1 binding sites, among them FOXO1, p21 and p57. In this system Runx1 binds and down-regulates FOXO1 and p21 (Wang et al., 2011 and Figure 1-2). These and other genes show a heterogeneous expression pattern that is dependent on phosphorylated Runx1 (Wang et al., 2011). This evidence points out that post-translational modification is essential for Runx1 function in these cells, a connection that already has been made for hematopoiesis (most recently Yoshimi et al., 2012). Notably, the down-regulation of p21 by Runx1 is similar to

that observed in the skin (Hoi et al., 2010). Therefore one might expect MCF10A cells without Runx1 to fail to grow too, but this did not occur. Breast cancer cell lines continue to grow and in MCF10A cells Runx1 loss even leads to hyper-proliferation that is strictly dependent on FOXO1 up-regulation (Wang et al., 2011; Scheitz et al., 2012). Alternatives to FOXO1 compensation mechanism may be present for some human cancer cells that express high levels of Runx1, while some cells cannot grow in the absence of Runx1 despite high FOXO1 expression (Scheitz et al., 2012). Runx1 presence in triple negative cancers and its suppression of p21 suggest an oncogenic function, but its connection with FOXO1 is rather tumor-suppressor like. The data on Runx1 and p21 suggest that Runx1 is tightly interwoven in this network of cell cycle regulators that inhibit cyclin dependent kinases, which in return phosphorylate Runx1 at least in breast cells (Wang et al., 2011). This supports the notion that Runx1 is a complex regulator and there are various mechanisms by which it could act as a tumor promoter or suppressor, even in subclasses of the same cancer and it emphasizes that we still have a lot to learn about Runx1 in the different breast cancer types.

Finally, there is evidence for Runx1 involvement in endometroid carcinoma and to date the data suggest that similar to SCCs Runx1 function is conserved across subtypes. Similar to skin SCCs, Runx1 is expressed close to the invasive edge in endometroid carcinoma and, moreover, is required for myometrial invasion (Planagumà et al., 2004). During cancer development Runx1 but also p21 levels increase gradually, peaking at invasion (Planagumà et al., 2006). While the authors do not have evidence favoring a direct or indirect interaction, they

propose 3 hypotheses, all involving stabilization of p21 levels leading to suppressed cell proliferation while inducing differentiation. Notably, p21 is more broadly expressed than Runx1 and levels correlate with proliferation in cells lacking Runx1 (Planagumà et al., 2006). Hence, it is possible that in endometroid cancer Runx1 expression mediates a switch from a proliferative to an invasive phenotype, matching the Runx1 expression pattern peaking at myometrial invasion.

In addition to p21, Runx1 expression shows a correlation with MMP-2 and-9 (Planagumà et al., 2011 and Figure 1-2), two gelatinases that are key to degrading the basement membrane. In both endometroid and ovarian carcinoma the MMPs and Runx1 are expressed most strongly at the invasive edge during the stage of invasion. These data support a hypothesis in which Runx1 mediates myometrial invasion through blocking proliferation and inducing membrane degradation. Although Planagumà et al. (2011) do not analyze if Runx1 binds MMP-2 and -9 promoters directly, it has been shown previously that Runx1 and Runx2, bind to and cause promoter activation of MMP-9 in bone metastatic cancer cells (Pratap et al., 2005). They also show that Runx2 regulates bone cancer cell invasion, and its overexpression causes increased invasion potential in metastatic (MDA-MB-231) and non-metastatic (MCF-7) breast cancer cell lines. It is possible that Runx1 also mediates bone cancer invasion given the overlap of the evidence in endometroid cancer and the MMP-9 promoter binding bone cancer, a possibility that should be evaluated further.

Lastly, Planagumà et al. (2011) also uncover a strong correlation of Runx1 expression at the proliferative edge with the expression of transcription factor ETV5/ERM in both ovarian and endometroid cancer. ETV5/ERM directly binds the promoter and regulates MMP-2 expression (Monge et al., 2007). Hence it would be plausible that Runx1 activates MMP-9 and ETV5/ERM transcription, which in turn activates MMP-2 to lead to myometrial infiltration. Interestingly, expression of ETV5/ERM has also been connected to malignancy of esophageal SCC and other malignancies (Yuen et al., 2011).

It is striking how many cancers appear to be characterized by regulatory networks involving Runx1 and how fast this list has grown in the recent years. It remains to be seen, if more cancers will be added to this list and how Runx1 function in normal homeostasis relates to its role in cancer formation, progression, and metastasis for each tissue type.

1.7 Concluding remarks

The common theme emerging for Runx1 function across tissues is its implication in stem cell regulation and in cancer. In addition to hematopoietic stem cells and neural progenitors (Kanaykina et al., 2010) in the last few years Runx1 has been reportedly expressed in 3 other stem cell populations: hair follicle, oral epithelium, and potentially in the intestine. Strong functional and genetic studies implicate Runx1 as a key regulator of both HSCs and HFSCs function. It will be interesting to examine how global Runx1 function is as a regulator of other adult tissue stem cells and cancers. Association and functional preliminary data in

breast, colon, intestinal and prostate suggest compelling directions for future studies. Moreover, the link between immune diseases and malfunctions of Runx1 requires further in depth investigation.

Runx1 roles in HSCs and HFSCs overlap to some extent but are distinct. In both systems early Runx1 expressing cells mark SC precursors and are required for adult tissue maintenance. Additionally, Runx1 is also expressed in the short-lived progenitors that form the first or 'primitive' wave of hair and hematopoietic lineages, respectively. Next, Runx1 signaling is essential to control proper adult SC activation and differentiation and overall Runx1 loss leads to a lack of differentiated blood and hair lineages. Naturally, there are differences in the details of these processes. Most importantly, impediment of HFSC activation in adulthood and timely emergence of hair placodes (which include the precursors of adult HFSCs) due to Runx1 loss is reversible. However, hematopoietic SCs cannot recover from Runx1 loss, since fetuses die prematurely due to secondary effects (Chen et al., 2009). Moreover, loss of Runx1 in the environment did not have an effect on the proper emergence of adult HSCs (Chen et al., 2009), while the adult HFSCs showed severe defects in maturation that became manifested by adulthood (Osorio et al., 2010). By focusing our work on new tissues we can begin to understand the capacity of Runx1 as a potential conductor of the stem cell concert that is performed daily in our bodies.

While there seem to be many common themes of Runx1 function in tissue stem cells, the opposite seems to be true for cancer development. So far we see three independent mechanisms through which Runx1 can regulate cancer

development. First, Runx1 represses p21 expression and thus positively regulates cell cycle progression in several tissues. Additionally, Runx1 is required for Stat3 phosphorylation, supporting cellular proliferation. Lastly, Runx1 binds and activates MMPs to potentially facilitate invasion. Notably, Runx1 modulates Wnt signaling in normal skin biology and it is possible that this connection applies to cancer, since Wnt activation is a major path to cancer promotion (Polakis, 2012). Each of these mechanisms alone is associated with cancer formation and together they indicate that most frequently Runx1 acts as a tumor promoter in epithelial tissues. Whether these mechanisms are mutually exclusive or may occur in the same tumor type remains to be seen. Further research is required to connect the branches of this cancer-regulating network with Runx1 as a central player. However, we have to be careful not to limit ourselves to a tumor-promoting function, since compelling evidence in mice showed that Runx1 can act as a tumor suppressor in some epithelial tissues. By all means we need to keep an open mind in further exploring this remarkable protein as it turns out clearly that Runx1 is not a one trick pony!

Table 1-1: Summary of Runx1 studies in solid cancers and autoimmune diseases

	Disease Type	Study Type	Reference
Cancer	Skin SCC	• Conditional mouse KO model	Hoi et al., 2010
	Skin SCC	• Inducible mouse KO model • Tumor lineage tracing • Mouse and human *in vitro* culture • Human tissue staining and microarray meta-analysis	Scheitz et al., 2012
	Head and Neck SCC	• Human *in vitro* culture • Human tissue staining and microarray meta-analysis	Scheitz et al., 2012
	Endometroid carcinoma	• Human tumor tissue staining • Human tumor microarray data	Planagumà et al., 2004, 2006, 2011
	Triple Negative Breast Cancer	• Human *in vitro* culture • Human tumor microarray data	Wang et al., 2011
	PR+ Breast Cancer	• Human tumor genome sequencing	Banerji et al., 2012
	Prostate Cancer	• Human primary cancer vs. normal microarray	Yeh et al., 2009
	Prostate Cancer	• Human SNP association study	Huang et al., 2011
	Intestinal Cancer	• Inducible mouse KO model • Mouse KO vs WT tissue microarray	Fijneman et al., 2012
	Colon and Rectal Cancer	• Human SNP association study	Slattery et al., 2011
	Colorectal Cancer	• Human cancer vs. normal microarray meta-analysis	Pradhan et al., 2012

	Disease Type	Study Type	Reference
Autoimmune disease	Psoriasis	• Human SNP association study • Human tissue staining	Helms et al., 2003
	SLE	• Human SNP association study • Human tissue expression assays	Prokuina et al., 2002
	Rheumatoid arthritis	• Human SNP association study • Human tissue expression assays	Tokuhiro et al., 2003

REFERENCES

Alam M, Ratner D. 2001. Cutaneous Squamous Cell Carcinoma. *New England Journal of Medicine* **344**: 975–983.

Banerji S, Cibulskis K, Rangel-Escareno C, Brown KK, Carter SL, Frederick AM, Lawrence MS, Sivachenko AY, Sougnez C, Zou L, Cortes ML, Fernandez-Lopez JC, Peng S, Ardlie KG, Auclair D, Bautista-Piña V, Duke F, Francis J, Jung J, Maffuz-Aziz A, Onofrio RC, Parkin M, Pho NH, Quintanar-Jurado V, Ramos AH, Rebollar-Vega R, Rodriguez-Cuevas S, Romero-Cordoba SL, Schumacher SE, Stransky N, Thompson KM, Uribe-Figueroa L, Baselga J, Beroukhim R, Polyak K, Sgroi DC, Richardson AL, Jimenez-Sanchez G, Lander ES, Gabriel SB, Garraway L a., Golub TR, Melendez-Zajgla J, Toker A, Getz G, Hidalgo-Miranda A, Meyerson M. 2012. Sequence analysis of mutations and translocations across breast cancer subtypes. *Nature* **486**: 405–409.

Barker N, Ridgway R a, van Es JH, van de Wetering M, Begthel H, van den Born M, Danenberg E, Clarke AR, Sansom OJ, Clevers H. 2009. Crypt stem cells as the cells-of-origin of intestinal cancer. *Nature* **457**: 608–611.

Challen GA, Goodell MA. 2010. Runx1 isoforms show differential expression patterns during hematopoietic development but have similar functional effects in adult hematopoietic stem cells. *Experimental Hematology* **38**: 403–416.

Chen MJ, Yokomizo T, Zeigler BM, Dzierzak E, Speck NA. 2009. Runx1 is required for the endothelial to haematopoietic cell transition but not thereafter. *Nature* **457**: 887–891.

Fife BT, Pauken KE. 2011. The role of the PD-1 pathway in autoimmunity and peripheral tolerance. *Annals of the New York Academy of Sciences* **1217**: 45–59.

Fijneman RJA, Anderson RA, Richards E, Liu J, Tijssen M, Meijer GA, Anderson J, Rod A, O'Sullivan MG, Scott PM, Cormier RT. 2012. Runx1 is a tumor suppressor gene in the mouse gastrointestinal tract. *Cancer Science* **103**:593-599.

Friedman AD. 2009. Cell cycle and developmental control of hematopoiesis by Runx1. *Journal of Cellular Physiology* **219**: 520–524.

Ghozi MC, Bernstein Y, Negreanu V, Levanon D, Groner Y. 1996. Expression of the human acute myeloid leukemia gene AML1 is regulated by two promoter

regions. *Proceedings of the National Academy of Sciences of the United States of America* **93**: 1935–1940.

Griffi CEM, Barker JNWN. 2007. Psoriasis 1 Pathogenesis and clinical features of psoriasis. Lancet 370: 264–271.

Helms C, Cao L, Krueger JG, Wijsman EM, Chamian F, Gordon D, Heffernan M, Daw JAW, Robarge J, Ott J, Kwok P-Y, Menter A, Bowcock AM. 2003. A putative RUNX1 binding site variant between SLC9A3R1 and NAT9 is associated with susceptibility to psoriasis. *Nature Genetics* **35**: 349–356.

Hoi C, Lee SE, Lu S-Y, McDermitt DJ, Osorio KM, Piskun C, Peters R, Paus R, Tumbar T. 2010. Runx1 directly promotes proliferation of hair follicle stem cells and epithelial tumor formation in mouse skin. *Molecular and Cellular Biology* **30**: 2518–2536.

Huang S-P, Lan Y-H, Lu T-L, Pao J-B, Chang T-Y, Lee H-Z, Yang W-H, Hsieh C-J, Chen L-M, Huang L-C, Ting W-C, Bao B-Y. 2011. Clinical significance of runt-related transcription factor 1 polymorphism in prostate cancer. *BJU International* **107**: 486–492.

Ishimatsu-Tsuji Y, Moro O, Kishimoto J. 2005. Expression profiling and cellular localization of genes associated with the hair cycle induced by wax depilation. *The Journal of Investigative Dermatology* **125**: 410–420.

Kanaykina N, Abelson K, King D, Liakhovitskaia A, Schreiner S, Wegner M, Kozlova EN. 2010. In vitro and in vivo effects on neural crest stem cell differentiation by conditional activation of Runx1 short isoform and its effect on neuropathic pain behavior. *Upsala Journal of Medical Sciences* **115**: 56–64.

Kurokawa M, Tanaka T, Tanaka K, Hirano N, Ogawa S, Mitani K, Yazaki Y, Hirai H. 1996. A conserved cysteine residue in the runt homology domain of AML1 is required for the DNA binding ability and the transforming activity on fibroblasts. *The Journal of Biological Chemistry* **271**: 16870–16876.

Lapouge G, Youssef KK, Vokaer B, Achouri Y, Michaux C, Sotiropoulou P a, Blanpain C. 2011. Identifying the cellular origin of squamous skin tumors. *Proceedings of the National Academy of Sciences of the United States of America* **108**: 7431–7436.

Lee J, Tumbar T. 2012. Hairy tale of signaling in hair follicle development and cycling. *Seminars in Cell & Developmental Biology*.

Levanon D, Brenner O, Negreanu V, Bettoun D, Woolf E, Eilam R, Lotem J, Gat U, Otto F, Speck N, Groner Y. 2001. Spatial and temporal expression pattern of Runx3 (Aml2) and Runx1 (Aml1) indicates non-redundant functions during mouse embryogenesis. *Mechanisms of Development* **109**: 413–417.

Mangan JK, Speck NA. 2011. RUNX1 mutations in clonal myeloid disorders: from conventional cytogenetics to next generation sequencing, a story 40 years in the making. *Critical Reviews in Oncogenesis* **16**: 77–91.

Merrill B, Gat U, DasGupta R, Fuchs E. 2001. Tcf3 and Lef1 regulate lineage differentiation of multipotent stem cells in skin. *Genes & Development* **15**: 1688–1705.

Miyoshi H, Shimizu K, Kozu T, Maseki N, Kaneko Y, Ohki M. 1991. t(8;21) breakpoints on chromosome 21 in acute myeloid leukemia are clustered within a limited region of a single gene, AML1. *Proceedings of the National Academy of Sciences of the United States of America* **88**: 10431–10434.

Monge M, Colas E, Doll A, Gonzalez M, Gil-Moreno A, Planaguma J, Quiles M, Arbos MA, Garcia A, Castellvi J, Llaurado M, Rigau M, Alazzouzi H, Xercavins J, Alameda F, Reventos J, Abal M. 2007. ERM/ETV5 up-regulation plays a role during myometrial infiltration through matrix metalloproteinase-2 activation in endometrial cancer. *Cancer Research* **67**: 6753–6759.

Morris RJ, Liu Y, Marles L, Yang Z, Trempus C, Li S, Lin JS, Sawicki J a, Cotsarelis G. 2004. Capturing and profiling adult hair follicle stem cells. *Nature Biotechnology* **22**: 411–417.

Ortt K, Raveh E, Gat U, Sinha S. 2008. A Chromatin Immunoprecipitation Screen in Mouse Keratinocytes Reveals Runx1 as a Direct Transcriptional Target Of DNp63. *Journal of Cellular Biochemistry* **104**: 1204–1219.

Osorio KM, Lee SE, McDermitt DJ, Waghmare SK, Zhang YV, Woo HN, Tumbar T. 2008. Runx1 modulates developmental, but not injury-driven, hair follicle stem cell activation. *Development (Cambridge, England)* **135**: 1059–1068.

Osorio KM, Lilja KC, Tumbar T. 2011. Runx1 modulates adult hair follicle stem cell emergence and maintenance from distinct embryonic skin compartments. *The Journal of Cell Biology* **193**: 235–250.

Planagumà J, Díaz-Fuertes M, Gil-Moreno A, Abal M, Monge M, García A, Baró T, Thomson TM, Xercavins J, Alameda F, Reventós J. 2004. A differential gene expression profile reveals overexpression of RUNX1/AML1 in invasive endometrioid carcinoma. *Cancer Research* **64**: 8846–8853.

Planagumà J, Gonzalez M, Doll A, Monge M, Gil-Moreno A, Baró T, García A, Xercavins J, Alameda F, Abal M, Reventós J. 2006. The up-regulation profiles of p21WAF1/CIP1 and RUNX1/AML1 correlate with myometrial infiltration in endometrioid endometrial carcinoma. *Human Pathology* **37**: 1050–1057.

Planagumà J, Liljeström M, Alameda F, Bützow R, Virtanen I, Reventós J, Hukkanen M. 2011. Matrix metalloproteinase-2 and matrix metalloproteinase-9 codistribute with transcription factors RUNX1/AML1 and ETV5/ERM at the invasive front of endometrial and ovarian carcinoma*. Human Pathology* **42**: 57–67.

Polakis P. 2012. Wnt signaling in cancer. *Cold Spring Harbor Perspectives in Biology* **4**.

Pradhan MP, Prasad NK, Palakal MJ. 2012. A systems biology approach to the global analysis of transcription factors in colorectal cancer. *BMC Cancer* **12**: 331.

Pratap J, Javed A, Languino LR, Andre J, Wijnen V, Stein JL, Stein GS, Lian JB. 2005. The Runx2 Osteogenic Transcription Factor Regulates Matrix Metalloproteinase 9 in Bone Metastatic Cancer Cells and Controls Cell Invasion. *Molecular and Cellular Biology* **25**: 8581–8591.

Prokunina L, Castillejo-López C, Oberg F, Gunnarsson I, Berg L, Magnusson V, Brookes AJ, Tentler D, Kristjansdóttir H, Gröndal G, Bolstad AI, Svenungsson E, Lundberg I, Sturfelt G, Jönssen A, Truedsson L, Lima G, Alcocer-Varela J, Jonsson R, Gyllensten UB, Harley JB, Alarcón-Segovia D, Steinsson K, Alarcón-Riquelme ME. 2002. A regulatory polymorphism in PDCD1 is associated with susceptibility to systemic lupus erythematosus in humans. *Nature Genetics* **32**: 666–669.

Raveh E, Cohen S, Levanon D, Negreanu V, Groner Y, Gat U. 2006. Dynamic expression of Runx1 in skin affects hair structure. *Mechanisms of Development* **123**: 842–850.

Sano S, Itami S, Takeda K, Tarutani M, Yamaguchi Y, Miura H, Yoshikawa K, Akira S, Takeda J. 1999. Keratinocyte-specific ablation of Stat3 exhibits impaired skin remodeling, but does not affect skin morphogenesis. *The EMBO Journal* **18**: 4657–4668.

Sano S, Kira M, Takagi S, Yoshikawa K, Takeda J, Itami S. 2000. Two distinct signaling pathways in hair cycle induction: Stat3-dependent and -

independent pathways. *Proceedings of the National Academy of Sciences of the United States of America* **97**: 13824–13829.

Scheitz CJF, Lee TS, McDermitt D, Tumbar T. 2012. Defining a tissue stem cell-driven Runx1/Stat3 signalling axis in epithelial cancer. *The EMBO Journal*.

Schneider MR, Schmidt-Ullrich R, Paus R. 2009. The hair follicle as a dynamic miniorgan. *Current Biology* : **19**: 132-142.

Slattery ML, Lundgreen A, Herrick JS, Caan BJ, Potter JD, Wolff RK. 2011. Associations between genetic variation in RUNX1, RUNX2, RUNX3, MAPK1 and eIF4E and risk of colon and rectal cancer: additional support for a TGF-β-signaling pathway. *Carcinogenesis* **32**: 318–326.

Soma T, Ishimatsu-Tsuji Y, Tajima M, Kishimoto J. 2006. Runx1 transcription factor is involved in the regulation of KAP5 gene expression in human hair follicles. *Journal of Dermatological Science* **41**: 221–224.

Tokuhiro S, Yamada R, Chang X, Suzuki A, Kochi Y, Sawada T, Suzuki M, Nagasaki M, Ohtsuki M, Ono M, Furukawa H, Nagashima M, Yoshino S, Mabuchi A, Sekine A, Saito S, Takahashi A, Tsunoda T, Nakamura Y, Yamamoto K. 2003. An intronic SNP in a RUNX1 binding site of SLC22A4, encoding an organic cation transporter, is associated with rheumatoid arthritis. *Nature Genetics* **35**: 341–348.

Tumbar T, Guasch G, Greco V, Blanpain C, Lowry WE, Rendl M, Fuchs E. 2004. Defining the epithelial stem cell niche in skin. *Science (New York, N.Y.)* **303**: 359–363.

Waghmare SK, Bansal R, Lee J, Zhang YV, McDermitt DJ, Tumbar T. 2008. Quantitative proliferation dynamics and random chromosome segregation of hair follicle stem cells. *The EMBO Journal* **27**: 1309–1320.

Wang L, Brugge JS, Janes KA. 2011. Intersection of FOXO- and RUNX1-mediated gene expression programs in single breast epithelial cells during morphogenesis and tumor progression. *Proceedings of the National Academy of Sciences of the United States of America* **108**: 803–812.

White AC, Tran K, Khuu J, Dang C, Cui Y, Binder SW, Lowry WE. 2011. Defining the origins of Ras/p53-mediated squamous cell carcinoma. *Proceedings of the National Academy of Sciences of the United States of America* **108**: 7425–7430.

Yan J, Liu Y, Lukasik SM, Speck N a, Bushweller JH. 2004. CBFbeta allosterically regulates the Runx1 Runt domain via a dynamic conformational equilibrium. *Nature Structural & Molecular Biology* **11**: 901–906.

Yeh H-Y, Cheng S-W, Lin Y-C, Yeh C-Y, Lin S-F, Soo V-W. 2009. Identifying significant genetic regulatory networks in the prostate cancer from microarray data based on transcription factor analysis and conditional independency. *BMC Medical Genomics* **2**: 70.

Yoshimi M, Goyama S, Kawazu M, Nakagawa M, Ichikawa M, Imai Y, Kumano K, Asai T, Mulloy JC, Kraft AS, Takahashi T, Shirafuji N, Kurokawa M. 2012. Multiple phosphorylation sites are important for RUNX1 activity in early hematopoiesis and T-cell differentiation. *European Journal of Immunology* **42**: 1044–1050.

Yuen H-F, McCrudden CM, Chan K-K, Chan Y-P, Wong ML-Y, Chan KY-K, Khoo U-S, Law S, Srivastava G, Lappin TR, Chan K-W, El-Tanani M. 2011. The role of Pea3 group transcription factors in esophageal squamous cell carcinoma. *The American Journal of Pathology* **179**: 992–1003.

Zhang YV, Cheong J, Ciapurin N, McDermitt DJ, Tumbar T. 2009. Distinct Self-Renewal and Differentiation Phases in the Niche of Infrequently Dividing Hair Follicle Stem Cells. *Cell Stem Cell* **5**: 267–278.

CHAPTER 2[2]

DEFINING A TISSUE STEM CELL DRIVEN RUNX1/STAT3 SIGNALING AXIS IN EPITHELIAL CANCER

2.1 Overview

Cancers and tissue stem cells (SCs) share similar molecular pathways for their self- renewal and differentiation. The race is on to identify unique pathways to specifically target the cancer, while sparing normal SCs. Here we uncover the transcription factor Runx1/AML1, a known hematopoietic and leukemia factor, albeit dispensable for normal adult SC homeostasis, as being important for some mouse and human epithelial cancers. We implicate Runx1 as a SC- intrinsic gene in mouse hair follicle and oral epithelia by genetic lineage tracing in adulthood. Runx1- expressing SCs, but not other cells that ectopically up-regulate Runx1 by injury and inflammation, are at the skin tumor origin. Runx1 loss impairs tumor initiation and maintenance and the growth of oral, skin, and ovarian epithelial human cancer cells. Runx1 stimulates Stat3 signaling via direct transcriptional repression of SOCS3 and SOCS4 and this is essential for cancer cell growth. Thus, Runx1 is a broader epithelial SC and cancer factor than previously recognized,

[2] This work has been published "Scheitz CJF, Lee TS, McDermitt DJ & Tumbar T (2012) Defining a tissue stem cell-driven Runx1/Stat3 signalling axis in epithelial cancer. EMBO doi: 10.1038/emboj.2012.270" and is reprinted here with permission. The author contributions are as follows: David McDermitt performed and evaluated hair follicle stem cell homeostasis lineage tracing; Tae Lee executed all skin tumor lineage tracing experiments and significantly contributed to monitoring the $Kras^{G12D}$ mice. Cornelia Scheitz designed the project, performed, analyzed and interpreted all other experiments created all figures and wrote the manuscript. Tudorita Tumbar designed the project, analyzed and interpreted data and wrote the manuscript.

and qualifies as an attractive potential target for both prevention and therapy of several epithelial cancers.

2.2 Introduction

Increasing evidence, including genetic lineage tracing, seems to support the theory that some cancers, such as skin basal and squamous cell carcinoma (Youssef et al., 2010; Lapouge et al., 2011; White et al., 2011), intestinal cancer, adenocarcinoma, and leukemia originate in tissue stem cells (SCs) (Visvader, 2011). Alternatively, some cancers might arise through de-differentiation of cells that hijack the normal SC self-renewal pathways (Visvader, 2011). In either case, for cancer therapies to work we must identify factors that are essential for the survival and proliferation of cancer cells, yet dispensable for normal SCs. There could be selective pressure among cancer cells to preserve the genomic loci of factors essential for their own growth and survival, and thus these factors might not surface in conventional human genetic mutation screens. In fact, such factors could be difficult to identify without in depth knowledge about the genetic control of normal and cancer SCs, which are difficult to access directly in humans.

We recently identified Runx1 as a factor involved, but dispensable for mouse skin epithelial hair follicle stem cells (HFSCs) proliferation in vivo (Osorio et al., 2008). In contrast, Runx1 was strictly required for mouse skin papilloma and squamous cell carcinoma (SCC) formation and for cultured keratinocyte proliferation (Hoi et al., 2010). This makes Runx1 an attractive potential target for cancer that would not affect normal tissue homeostasis. However, the

significance of our mouse finding to human cancer biology remained unclear. Human skin SCC is the second most common cancer after skin basal cell carcinoma (BCC), with ~0.8 million new cases per year in the USA alone, of which ~8% are metastatic and cannot be treated with current methods (Alam and Ratner, 2001). Head and neck (oral) SCC are extremely invasive and currently un-curable, and have several regulatory pathways in common with skin SCC (Molinolo et al., 2009).

Runx1 is recognized as a human leukemia and hematopoietic stem cell factor (Mangan and Speck, 2011), but its role in human SCC and potentially other epithelial cancers is poorly understood. Recently Runx1 SNPs were associated with human colon and rectal cancer (Slattery et al., 2011) as well as prostate cancer (Yeh et al., 2009; Huang et al., 2011), and Runx1 was found expressed in endometrial and ovarian cancer (Planagumà et al., 2006, 2011). Moreover, triple negative (-Her2,-ER,-PR) human breast tumor growth in culture depends partially on Runx1 expression (Wang et al., 2011).

This evidence began to suggest that the leukemia factor Runx1 might be more broadly implicated in cancer than previously recognized. Here, we present a breadth of evidence that Runx1 is over-expressed in a significant fraction of human epithelial cancers, and is essential for the growth and survival of three of them: skin SCC, oral SCC, and in part for ovarian cancer. Our lineage tracing data in adult mice show for the first time by lineage tracing that Runx1 is expressed in adult HFSCs and in oral epithelium SCs. The oral epithelium SCs were previously

poorly characterized. Runx1 appears essential for solid tumor initiation and maintenance by upstream stimulation of Stat3 signaling, a central cancer pathway implicated in several epithelial cancers and in leukemia (Chan et al., 2004a, 2004b, 2008; Kim et al., 2009; Aziz et al., 2007; Yu and Jove, 2004; Li et al., 2011). Together with the notion that Runx1 is not critical for normal adult tissue function, our work places this gene as an important candidate for epithelial cancer targeting.

2.3 Results

2.3.1 Runx1 is over-expressed in various epithelial tumors

Previously, we found Runx1 expressed in the same pattern in normal mouse and human skin (Hoi et al., 2010). Runx1 has a dynamic expression in the hair follicle during its homeostatic cycle (hair cycle), which consists of the morphologically distinct stages of growth and proliferation (anagen), regression (catagen) and rest (telogen) (Figure 2-1) (Tumbar, 2012). During quiescence (telogen) Runx1 marks a subset of the hair follicle bulge and hair germ cells where the SCs reside (Osorio et al., 2008). Moreover, Runx1 protein is detected in the bulge during anagen when HFSCs self-renew (Osorio et al., 2008, Zhang et al., 2009). The embryonic precursors of HFSCs express Runx1 (Osorio et al., 2011). Lineage tracing to formally prove Runx1 expression in adult HFSCs has not yet been done, but loss of Runx1 at different stages of the hair cycle delays HFSC activation from quiescence (Osorio et al., 2008), and slows down bulge cell proliferation during anagen (Hoi et al., 2010). Finally, we found Runx1 expressed at high levels in

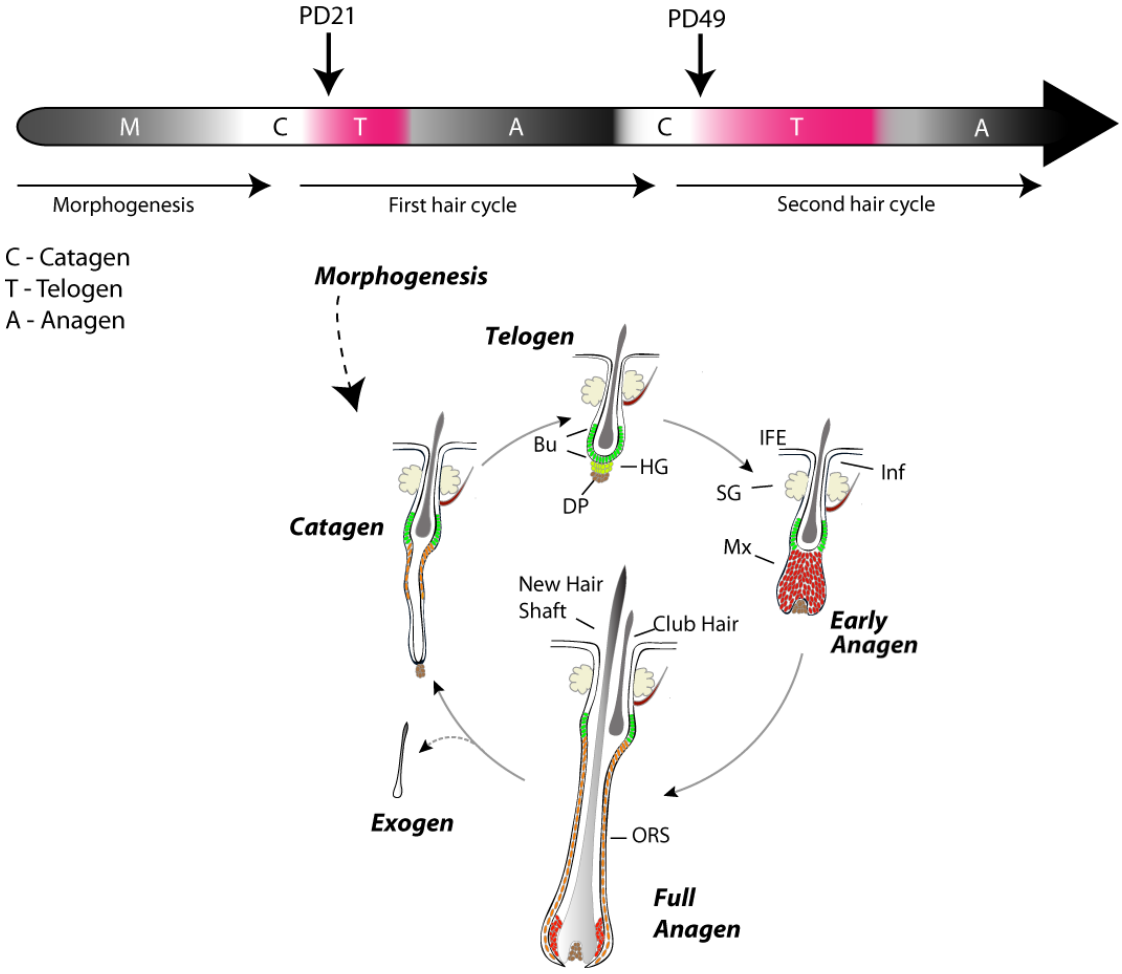


Figure 2-1: Hair Cycle

Top: Scheme of hair cycle stages and their progression over time. Skin color observed in BL6 mice is indicated at each stage: pink for telogen, white for catagen, black for anagen. Bottom: Diagram of hair morphology and structure remodeling at different hair cycle stages. IFE = interfollicular epidermis; Bu = bulge; HG = hair germ; DP = dermal papilla; Inf= infundibulum ; SG = sebaceous gland; Mx = matrix; ORS = outer rooth sheath.

mouse skin papilloma and SCC, and Runx1 deficiency impaired mouse skin tumorigenesis (Hoi et al., 2010).To examine the expression pattern of Runx1 in human epithelial cancers we first surveyed the Oncomine public database, which contains microarray data comparing primary tumors and normal tissue. We found Runx1 to be over-expressed in several types of cancers in 47 studies and under-expressed in only 5 out of 138 studies analyzed. The data are summarized in Figure 2-2A and Table 2-1. Notably, comparing tumor to normal tissue Runx1 was among the top 1% of all over-expressed genes in head and neck, and skin SCC, and was highly expressed in several other epithelial tumors such as esophageal SCC, cervical carcinoma, and colon adenocarcinoma. Next, we verified Runx1 expression by western blot and immuno-fluorescence staining of a panel of human cancer cell lines from: breast (MCF-7; MDA-MB-231), prostate (PC-3; DU-145), colon (HCT116; CCSC, Sikandar et al., 2010), ovary (OVCAR-3; SKOV3), head and neck (oral) SCC (SCC66, SCC74, SCC125, White et al., 2007), and skin SCC (SCC13, Rheinwald and Beckett, 1980, 1981), as well as a mouse skin SCC cell line (Pam212, Yuspa et al., 1980). All cancer cell lines tested expressed more Runx1 relative to normal mouse keratinocytes, while we detected close to no Runx1 expression in mouse fibroblast, which served as negative control (Figure 2-2B-D, Figure 2-3A, B). Attesting to the specificity of our antibody, we detected no Runx1 signal by immuno-staining of knockout mouse keratinocytes (not shown). The different cancer cell lines showed variable levels and some contained only a sub-fraction of Runx1-positive cells. Notably, all head and neck and skin SCC cell lines showed a consistently high level of Runx1 expression.

Figure 2-2: Runx1 in human cancer

A: Meta-analysis of microarray studies comparing Runx1 levels in tumor vs. normal tissue. Increased Runx1 expression (fold-change) is marked red, whereas a decrease is blue. B: Western blot of mouse keratinocytes and mouse and human SCC skin (13) and head and neck (125,66, and 74) cells for Runx1 and actin. C: Summary of Runx1 expression in cancer cell lines of various origins derived from western blot (B) and stainings (D). D: Samples images of Runx1 staining in human tumor cell lines. E: Sample panel of mouse and human skin tumors stained for Runx1 (red), Keratin 5,8,15, and Ki67 (green) and DNA (blue). Dashed line marks the interface of stroma (*) and tumor. A human HF at equal exposure is presented for comparison. Image with increased brightness of Runx1 illustrates that Runx1 is present in this HF.

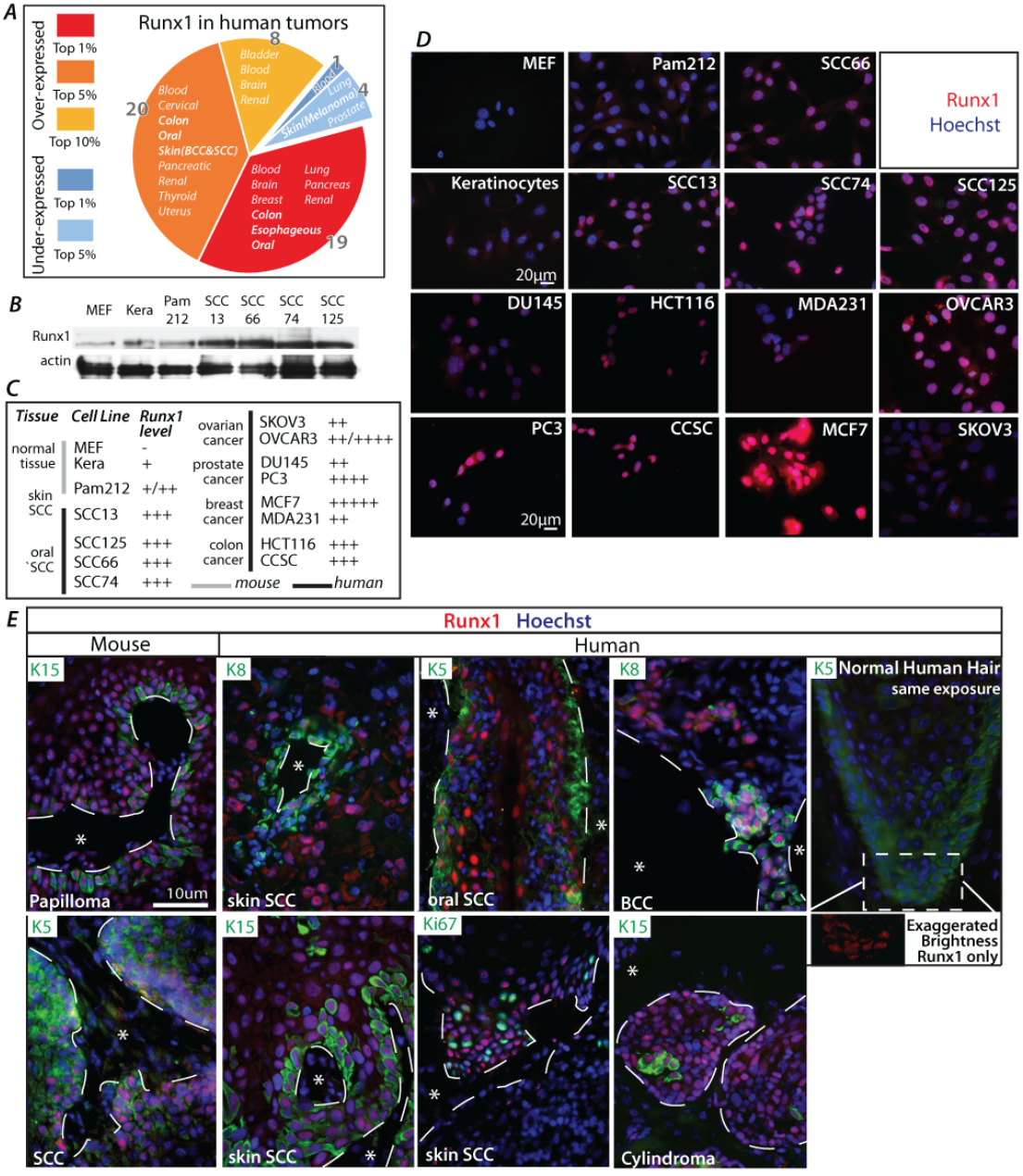

A
Runx1 in human tumors
Over-expressed
Top 1%
Top 5%
Top 10%
Under-expressed
Top 1%
Top 5%
8
Bladder
Blood
Brain
Renal
1
Blood
4
Lung
Skin(Melanoma)
Prostate
20
Blood
Cervical
Colon
Oral
Skin(BCC&SCC)
Pancreatic
Renal
Thyroid
Uterus
Blood
Brain
Breast
Colon
Esophageous
Oral
Lung
Pancreas
Renal
19
B
MEF
Kera
Pam 212
SCC 13
SCC 66
SCC 74
SCC 125
Runx1
actin
C
Tissue
Cell Line
Runx1 level
normal tissue
MEF -
Kera +
skin SCC
Pam212 +/++
oral SCC
SCC13 +++
SCC125 +++
SCC66 +++
SCC74 +++
ovarian cancer
SKOV3 ++
OVCAR3 ++/++++
prostate cancer
DU145 ++
PC3 ++++
breast cancer
MCF7 +++++
MDA231 ++
colon cancer
HCT116 +++
CCSC +++
mouse
human
D
MEF
Pam212
SCC66
Runx1
Hoechst
Keratinocytes
SCC13
SCC74
SCC125
20μm
DU145
HCT116
MDA231
OVCAR3
PC3
CCSC
MCF7
SKOV3
20μm
E
Runx1 Hoechst
Mouse
Human
K15
Papilloma
10um
K8
skin SCC
K5
oral SCC
K8
BCC
K5
Normal Human Hair
same exposure
Exaggerated Brightness Runx1 only
K5
SCC
K15
skin SCC
Ki67
skin SCC
K15
Cylindroma

Table 2-1: Meta-analysis of tumor vs. normal tissue microarray studies in Oncomine.

Study	Organ	#	fold-change		Rank
Stegmaier	Blood[1]	9	4.9	*	Top 1%
Haferlach	Blood[1]	174	2.5	***	
Zhan	Blood[2]	44	4.1	**	
Choi	Blood[1]	19	-2.1	*	
Pyeon	Cervical[3]	20	3.2	***	
TCGA	Colon[4]	22	5.9	***	
TCGA	Colon[4]	6	4.0	**	
Pyeon	Oral[3]	6	2.4	**	
FriersonHF	Oral[3]	16	180.0	**	
Badea	Pancreas[4]	39	5.0	***	
Logsdon	Pancreas[4]	10	222.6	**	
Yusenko	Renal[3]	26	6.6	***	
Riker	Skin (BCC)[3]	15	4.6	**	
Riker	Skin (SCC)[3]	11	2.4	*	Top 10%
Detwiller	Soft-tissue[6]	3	6.4	*	
Giordano	Thyroid[3]	10	2.3	**	
Giordano	Thyroid[3]	26	2.3	*	
Quade	Uterus[6]	4	26.7	*	
Quade	Uterus[6]	9	14.7	*	
Crabtree	Uterus[8]	50	2.4	**	
Haferlach	Blood[1]	147	2.3	***	
Haferlach	Blood[1]	542	2.1	***	
Haferlach	Blood[1]	79	2.6	***	
Lee	Brain[7]	22	3.9	*	
TCGA	Breast[3]	76	2.1	***	
TCGA	Breast[3]	36	2.0	***	
TCGA	Colon[4]	60	5.3	***	
TCGA	Colon[4]	101	4.8	***	
TCGA	Colon[4]	22	2.1	**	
Hong	Colon[3]	70	2.1	***	
Korkola	Embryo[3]	15	2.3	**	
Korkola	Teratoma	14	3.1	**	
Hao	Esophagus[4]	5	8.2	*	
Hao	Esophagus[3]	14	3.3	*	
Su	Lung[4]	27	2.3	*	
Bhattacharjee	Lung[3]	6	-17.9	*	
Ginos	Oral[3]	41	8.7	***	
Segara	Pancreas[3]	11	3.4	*	
Tomlins	Prostate[9]	13	-2.4	*	
Yusenko	Renal[3]	19	2.9	*	
Gumz	Renal[3]	26	6.5	*	
Talantov	Skin[5]	45	-4.6	**	
Talantov	Skin[5]	18	-5.8	*	
Detwiller	Soft-tissue[6]	9	8.7	*	
Sanchez-Carbayo	Bladder[3]	81	2.8	**	Top 5%
Andersson	Blood[1]	11	2.7	*	
Haferlach	Blood[1]	70	2.2	***	
Haferlach	Blood1	23	2.6	**	
Sun	Brain[7]	22	4.4	***	
Bredel	Brain[7]	22	7.1	*	
Beroukhim	Renal[3]	32	3.5	**	
Beroukhim	Renal[3]	27	3.1	*	

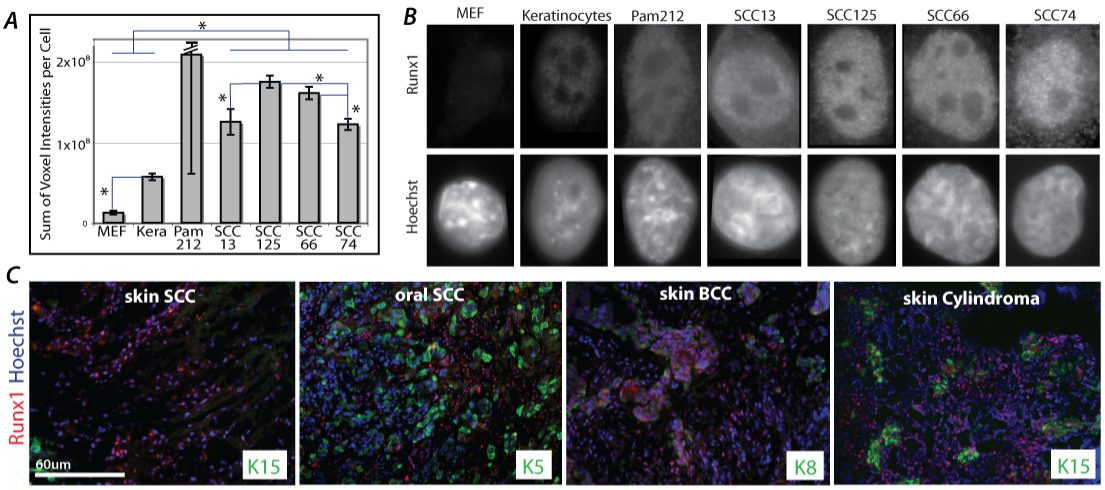


Figure 2-3: Runx1 in human cancers 2

A,B: Quantification (A) and images (B) of Runx1 staining in mouse keratinocytes and mouse and human SCC skin (13) and head and neck (125,66, and 74) cells. Staining was quantified cell wise. C: Runx1 staining in additional human tumor sections.

Finally, we checked for Runx1 expression in human primary tumors of skin cylindroma, skin SCCs and BCCs, as well as head and neck (oral) SCCs by immuno-fluorescence staining. We found Runx1 expressed in a subset of tumor cells including the proliferative edge, and these cells co-localized in part with Ki67 (a proliferation marker), K15 (a HFSC marker) or K8 (a malignancy marker, Casanova et al., 2004). The results are detailed in Table 2-2, Figure 2-2E, Figure 2-3C.

The widespread expression of Runx1 in human epithelial cancers uncovered here, together with the strict requirement of Runx1 for the formation of mouse skin SCC (Hoi et al., 2010) warranted further examination of a potential role of Runx1 in human solid tumors and prompted us to delve deeper in the essentials of its mechanism. Specifically we asked whether Runx1 expression is: (a) required at the tumor initiation, promotion, or maintenance stage; (b) found in normal SCs and in cells of tumor origin; (c) ectopically raised by injury and inflammation in cells of tumor origin; (d) essential for human epithelial cancer cell growth and survival; (e) coupled with known signaling pathways of epithelial cancers.

2.3.2 Runx1 is required for tumor initiation but dispensable for promotion

Previously, we showed that conditional deletion of Runx1 in mouse epidermis using a constitutively active K14-Cre allele drastically reduced papilloma and skin SCC formation (Hoi *et al.*, 2010). If Runx1 is to be used as a cancer target it is

Table 2-2: Runx1 expression in various human and mouse skin and oral cancer samples *in vivo*

	Human				*Mouse*
tumor type	***SCC***	***SCC***	***BCC***	***Cylindroma***	***SCC***
tumor tissue	***skin***	***oral***	***skin***	***skin***	***skin****
Source	Dr. J Vogel (NIH), Victorian Cancer Biobank	Dr. A Kobielak (USC)	Dr. J Vogel (NIH)	Dr. R Paus (Lübeck)	
# of samples with Runx1 / total # of samples	11/12	4/4	3/4	5/5	7/8
Runx1 expression pattern	a,b	a, some b	strong a, b	strong a and b	a, rare b
Other markers detected in tumors	K5, K15, K8, /, Ki67	K5, K15, K8, na, Ki67	K5, K15, K8, K10, Ki67	K5, K15, /, na, Ki67	K5, K15, K8, K10, Ki67

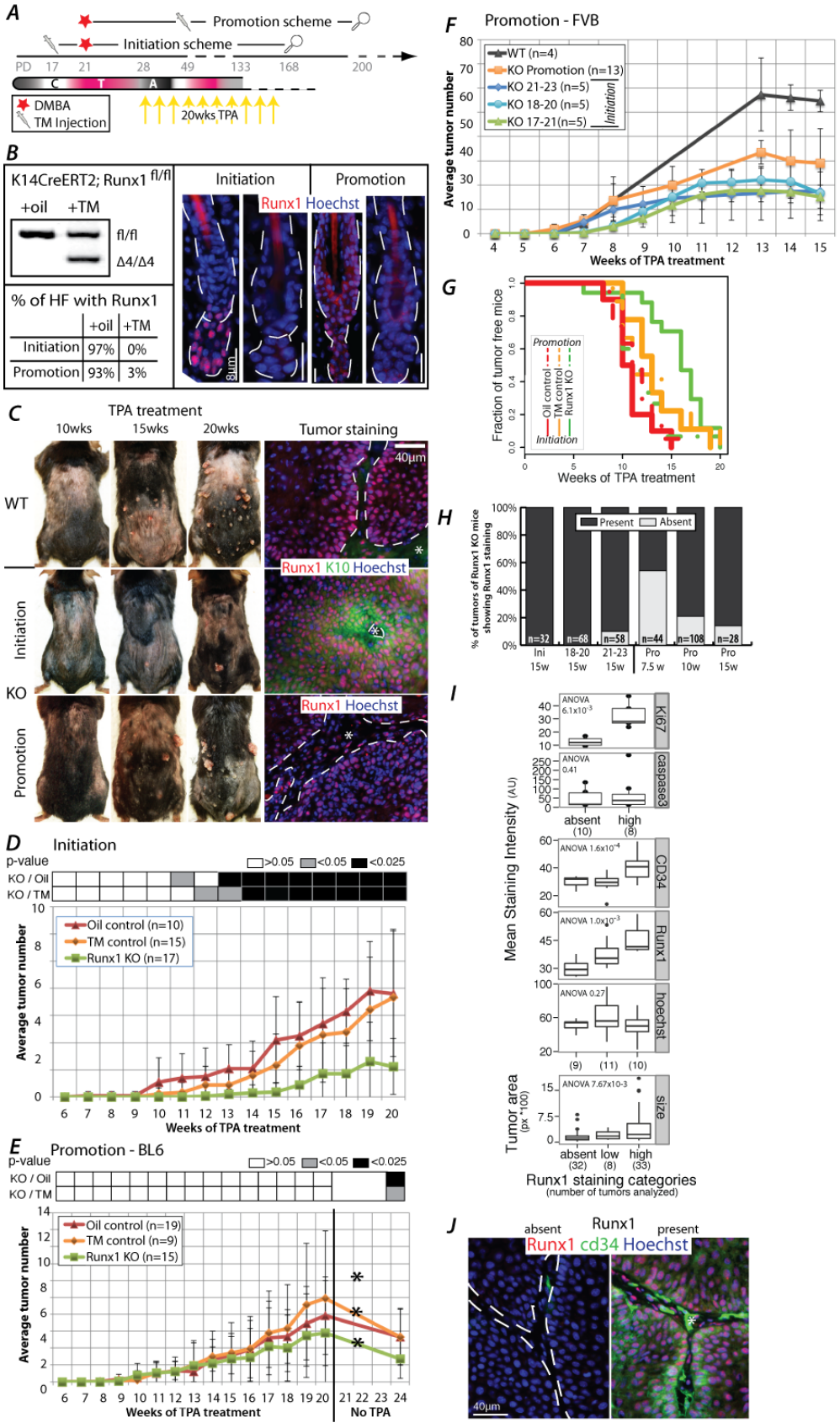
a = Leading edge of tumor; b = Disorganized tissue ; / = staining not detected; na = staining not available; * Hoi et al., 2010

important to determine the stage at which is acting in tumorigenesis (initiation, promotion, and maintenance). Thus, we deleted Runx1 by Tamoxifen (TM) injection in BL6 mice that carry the K14-CreERT2; $Runx1^{fl/fl}$ alleles at 3 different stages of tumorigenesis induced by a 2-step carcinogenesis protocol (Figure 2-4A, Figure 2-9A). The knockout (KO) targets the Runt domain in exon 4 rendering a null allele (Growney *et al.*, 2005). Four days after TM injection Runx1 KO was efficiently induced in over 90% of all HFs (Figure 2-4B). We applied 9,10-Dimethyl- 1,2-benzanthracene (DMBA) to induce point mutations in the DNA at 1^{st} telogen (PD21) and 12-O-tetradecanoylphorbol-13-acetate (TPA) for the following 20 weeks to promote the mutation by robust proliferation of skin epithelial cells (Abel *et al.*, 2009). As expected, after 20 weeks of TPA treatment, there were fewer tumors in the Runx1 KO mice TM-treated at PD17 relative to both control groups (TM no CreER and oil only) (Figure 2-4C top and middle,D,G). We refer to this treatment as the initiation scheme (Figure 2-4A). Tumors that did form in the KO mice expressed Runx1 at high levels, likely from the rare cells that escaped Cre activity, developed later on and were "young" in their development, as indicated by K10 expression (Figure 2-4C middle and Figure 2-5A, Santos *et al.*, 1997). These data from inducible KO mice with a BL6 background were in line with our previous results obtained from mice carrying the K14-Cre allele, which were a mix of CD1 and BL6 genetic backgrounds, activated in embryogenesis (Hoi *et al.*, 2010).

To differentiate between a role in tumor initiation and promotion, we injected

Figure 2-4: Runx1 KO at tumor initiation or promotion impairs SCC formation

A: Scheme of TM injection and DMBA/TPA treatments in mice to analyze Runx1 contribution to tumor initiation and promotion. Magnifying glass represents the end of monitoring mouse tumors. C, catagen, T, telogen, A, anagen. B: PCR of genomic DNA from mice of indicated genotype to detect the Runx1 excision band (Δ4) and quantification as well as representative images of Runx1 staining in HFs from both schemes in (A). C: Images of mice after 10,15,and 20 weeks of TPA treatment in either scheme and staining of frozen section of corresponding tumors. Dashed line marks the interface of stroma (*) and tumor. D: Tumor development of $Runx1^{fl/fl}$; K14-CreER mice injected with TM (green) or oil (red) and of $Runx1^{fl/fl}$ injected with TM (orange) treated as in the initiation scheme in (A). p-values obtained by pairwise t-test (average ± SD). E: Same as D but for the promotion scheme in (A). A star represents a significant drop of tumor numbers after TPA cessation. F: Tumor development in mixed FVB (average ± SD). G: Kaplan-Meier survival analysis for tumor onset in initiation (solid line, p-value = 7.52x10-6) and promotion schemes (dashed lines, p-value = 0.336) of BL6 mice. H: Quantification of tumor numbers sorted by Runx1 expression (where “Present” comprises “low” and “high” from I) in tumors derived from FVB/Bl6 mice. Ini = Initiation, Pro = Promotions, w= weeks. I: Quantification of size, CD34, Ki67, and caspase 3 in tumors grouped by their Runx1 levels. J: Staining of Runx1 and CD34 levels in FVB/BL6 derived promotion scheme tumors. Dashed line marks the interface of stroma (*) and tumor.


A
Promotion scheme
Initiation scheme
PD 17 21 28 49 133 168 200
C T A
DMBA
TM Injection
20wks TPA
B
K14CreERT2; Runx1 fl/fl
+oil +TM
fl/fl
Δ4/Δ4
% of HF with Runx1
+oil +TM
Initiation 97% 0%
Promotion 93% 3%
Initiation
Promotion
Runx1 Hoechst
50μm
C
TPA treatment
10wks 15wks 20wks
Tumor staining
40μm
WT
Initiation
KO
Promotion
Runx1 K10 Hoechst
Runx1 Hoechst
D
Initiation
p-value
>0.05 <0.05 <0.025
KO / Oil
KO / TM
Oil control (n=10)
TM control (n=15)
Runx1 KO (n=17)
Average tumor number
Weeks of TPA treatment
E
Promotion - BL6
p-value
>0.05 <0.05 <0.025
KO / Oil
KO / TM
Oil control (n=19)
TM control (n=9)
Runx1 KO (n=15)
Average tumor number
Weeks of TPA treatment
No TPA
F
Promotion - FVB
WT (n=4)
KO Promotion (n=13)
KO 21-23 (n=5)
KO 18-20 (n=5)
KO 17-21(n=5)
Initiation
Average tumor number
Weeks of TPA treatment
G
Fraction of tumor free mice
Promotion
Oil control
TM control
Runx1 KO
Initiation
Weeks of TPA treatment
H
% of tumors of Runx1 KO mice showing Runx1 staining
Present
Absent
n=32 n=68 n=58 n=44 n=108 n=28
Ini 15w
18-20 15w
21-23 15w
Pro 7.5 w
Pro 10w
Pro 15w
I
Mean Staining Intensity (AU)
Ki67
caspase3
CD34
Runx1
hoechst
Tumor area (px *100)
size
absent high
(10) (8)
absent low high
(32) (8) (33)
Runx1 staining categories
(number of tumors analyzed)
J
absent
Runx1
present
Runx1 cd34 Hoechst
40μm

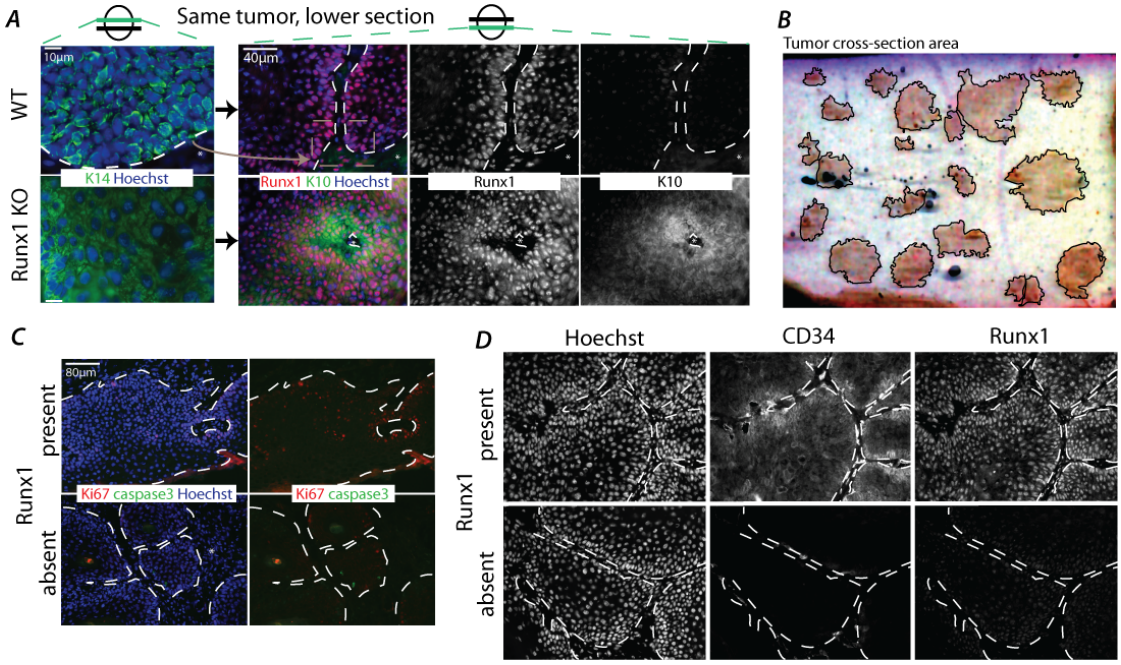


Figure 2-5: Runx1 expression affects expression of proliferation and stem cell markers

A: Sections from each a KO and WT mouse derived papilloma stained for K14 (upper section) and K10, Runx1 (lower section) from Figure 2-4C showing its channels separately for easier visibility. Dashed lines mark the interface of tumor and stroma (*). B: Scan of a representative tumor slides used for size approximation. D: Ki67 and caspase 3 staining in tumors lacking Runx1 or showing high levels. Quantification is shown in Figure 2-4J. E: Separate channels of CD34 staining from Figure 2-4I for better visibility of staining patterns.

TM 2.5-3 weeks after DMBA application when HF bulge cells divided at least once having passed through a full anagen (promotion scheme, Figure 2-4A,B). Surprisingly, Runx1 KO mice and control mice developed comparable numbers of tumors at a similar rate (Figure 2-4C bottom,E,G), indicating that Runx1 is dispensable for tumor promotion or growth at least in this genetic background. However, one month after the final TPA treatment tumor counts dropped by 53% in the Runx1 KO mice as compared to 33% in WT mice, resulting in a significant tumor difference between the KO and control animals(Figure 2-4E). Importantly, the only surviving tumors were those that still expressed Runx1 due to inefficient Cre activity (Figure 2-4C bottom).

To check the robustness of our phenotypes in a uniform genetic background highly prone to skin cancer (Abel *et al.*, 2009), we backcrossed our mice for 5 generations with Fvb background mice. As before, we found impaired tumorigenesis when Runx1 was deleted prior to tumor initiation (Figure 2-4F). This was true whether the Runx1 KO was induced 2 days before or 2 days after DMBA application either in telogen (PD18-20) or in anagen (PD21-23). When Runx1 was deleted 2.5 weeks after tumor initiation, 53% of tumors completely lacked Runx1 expression at week 7.5 of TPA treatment (Figure 2-4H,I). This was unlike our data from the initiation scheme, when no Runx1 negative tumors could be detected. This demonstrated that past the initiation stage tumors do not strictly require Runx1 for their growth under TPA treatment and was consistent with our data from the promotion scheme in the BL6 background. Suggestively,

these Runx1-negative tumors were smaller in size (Figure 2-4I bottom and Figure 2-5B) and their number decreased steadily from 7.5 to 15 weeks of TPA treatment (Figure 2-4H) indicating that tumors are short-lived in the absence of Runx1. Lower levels of Ki67 in these KO tumors support this loss, while apoptosis as assayed by Caspase 3 staining does not seem to play a significant role (Figure 2-4I and Figure 2-5C). Co-incidentally, the cancer SCs previously defined as a population of CD34- expressing papilloma cells (Malanchi *et al.*, 2008), were greatly diminished in tumors that lacked Runx1 expression (Figure 2-4I,J and Figure 2-5D). In general, cancer SCs are thought to maintain tumors and cause their re-occurrence for extended periods of time; CD34+ cancer SCs have been shown to initiate SCC formation upon serial transplantation for at least 3 generations (Malanchi *et al.*, 2008). Our finding suggests that Runx1 expression is important for the maintenance of the CD34+ tumor SCs. Moreover, Runx1 is only essential for tumor growth in the absence of a proliferative agent such as TPA, which can at least temporarily overcome Runx1 loss.

Overall, our experiments clearly distinguish an irreplaceable function of Runx1 in tumor initiation from a somewhat more context-dependent function in tumor promotion.

2.3.3 Runx1-expressing cells are hair follicle stem cells

We previously reported that Runx1 loss impaired adult HF bulge-cell proliferation (Osorio *et al.*, 2008; Hoi *et al.*, 2010), but it remained unclear whether this effect is SC-intrinsic or environmental, and if Runx1 is actually

expressed in the adult long-term HFSCs, although we did find it expressed in the embryonic precursors of HFSC (Osorio *et al.*, 2011). To address this question here we used lineage tracing involving an inducible CreER recombinase previously knocked into the Runx1 genomic locus (Samokhvalov *et al.*, 2007) and ROSA26R (Soriano, 1999) reporter mice. We injected TM at PD17 in these mice and sacrificed them at 1st telogen (PD21), 1st anagen (PD35), and 3rd anagen (~1-year old). In skin sections stained with X-Gal from the 1st telogen we observed labeling in ~50% of HFs marking ~5 cells per bulge (Figure 2-6A-C). Later on at the 1st and 3rd anagen we found prominent trails of X-Gal+ cells in the ORS (marked by K14) and the differentiated portion of the HF co-localizing with AE13 and AE15, as well as small clones of cells in the bulge (Figure 2-6B top, C-E). These X-Gal+ cells were completely absent in mice injected with oil-only (Figure 2-6B, bottom). These data demonstrate for the first time by lineage tracing that original Runx1-expressing adult bulge cells are long-term self-renewing and differentiating HFSCs.

2.3.4 Runx1-expressing HFSCs are at the origin of skin tumors

Since Runx1 is expressed in adult HFSCs, which can generate skin SCC, and Runx1 is important for tumor initiation it follows that it might be expressed in cells of tumor origin. However, epithelial skin tumors can originate in both HF and epidermis (Youssef *et al.*, 2010; Lapouge *et al.*, 2011; White *et al.*, 2011). Our data so far suggest that tumor cells might hijack the self-renewal/proliferation pathways of normal HFSCs to promote their growth. Since Runx1 expression is

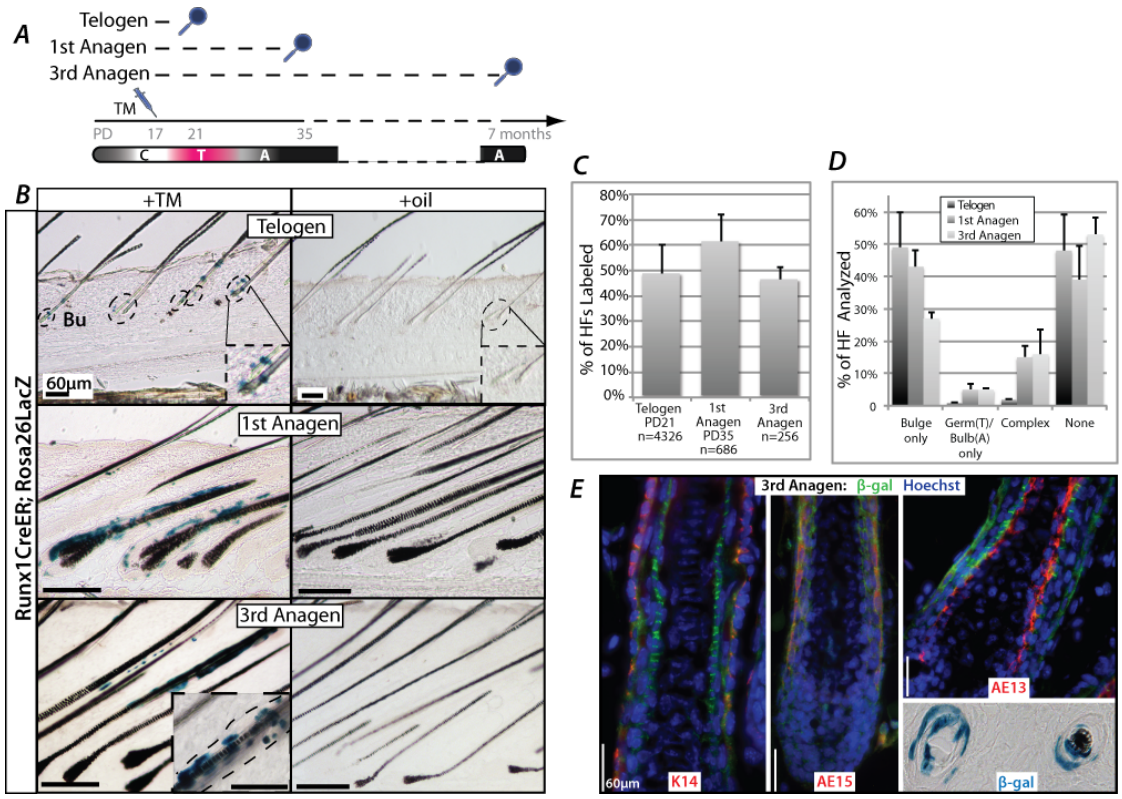


Figure 2-6: Lineage tracing shows Runx1 is expressed in HFSCs

A: Schemes to trace the contribution of Runx1 bulge cells to HF homeostasis. C, catagen; T, telogen; A, anagen. B: Representative images of the bulge lineage tracing (X-gal staining) 4 days after injection of oil or TM in Runx1-CreER xRosa26 mice sacrificed at Telogen and at 1st and 3rd Anagen. The inlay in 3rd Anagen TM image shows an additional high-magnification picture of an independent HF at that stage. C: Quantification of bulge and germ (Telogen) and bulb/bulge (Anagen) labeling (average ± SD). D: Quantification of lineages labeled after long-term chase in Telogen and 1st and 3rd Anagen (average ± SD). E: Staining of traced HFs at 3rd Anagen for the differentiation markers AE13 and AE15, and the outer root sheath marker K14. Bottom right highlights the X-gal staining of multiple layers in a cross-sectioned HF.

confined to the HF in normal skin, it was intriguing to find that all skin tumors in our initiation scheme expressed high levels of Runx1 (Hoi *et al.*, 2010). It is possible that some, but not all skin epithelial tumors generated in our protocol originated in HFSCs that normally expressed Runx1 and the remaining cells of tumor origin acquired ectopic Runx1 expression near the point of tumor initiation due to inflammation or injury induced by the treatment. We aimed to understand when Runx1 expression is aberrantly acquired in tumors and to what extent Runx1+ cells contribute to the cells of tumor origin. This would provide the context in which tumors are initiated and promoted.

The strategy was to induce genetic lineage tracing via TM injections in the Runx1-CreER; ROSA26R mice at different stages of tumorigenesis. We injected TM at PD17/18 in Runx1-CreER; ROSA26R mice and initiated the tumors with DMBA at PD21 (Figure 2-7A, Bulge scheme). This resulted in an average of 4.45 X-Gal+ cells/bulge in 80% of the HFs and 20% infundibular (Figure 2-7B,C, Bulge scheme) but no interfollicular epidermis labeling. After 20 weeks of TPA treatment, whole-mount X-Gal staining of the mouse back skin revealed 3/89 positive tumors in the TM-injected animals, as opposed to 0/71 in the oil controls (p-value: 0.06 Table 2-3). The positive tumors were completely blue suggesting monoclonal expansion, and expressed Runx1 (Figure 2-7D Bulge Scheme, Figure 2-8A). Of important note only ~2.3% of all bulge cells were X-Gal labeled upon TM induction due to low efficiency of the Runx1-CreER driven recombination. Using previously engineered Runx1-LacZ knock-in (KI) mice that report

Figure 2-7: Runx1 expressing HFSCs are at the origin of SCCs

A: Schemes of genetic lineage tracing in Runx1-CreER;Rosa26 mice to label and track bulge versus infundibular populations of Runx1+ cells contribution to tumorigenesis. Quantification of bulge (B) and infundibulum (C) labeling in mice treated as shown in each of the two schemes in (A). (average ± SD; n > 300HF). D: Whole mount (left panels) and skin sections of X-Gal positive and negative tumors from the schemes in (A). Counterstain is hematoxylin. E&F: Runx1-LacZ KI mice were subjected to a penetrating skin cut (A, red arrow) or TPA treatment (B, yellow arrows), sacrificed 4 days later, and skin sections analyzed by X-Gal staining. Representative images show increased Runx1 expression in the HF infundibulum upon either physical or oncogenic assault. In E and H an asterisk (*) indicate a normal skin area, a dash (#) the wound area. G&H: Runx1-CreER;Rosa26LacZ mice were subjected to TPA treatment (yellow arrows), followed by TM injection (syringe) to label the infundibulum. Bottom images show X-Gal staining of skin sections from mice after TM injection (G) or after TM injection and injury (H) at times indicated. Note lack of contribution from the infundibulum cells to either normal tissue homeostasis or injury repair.

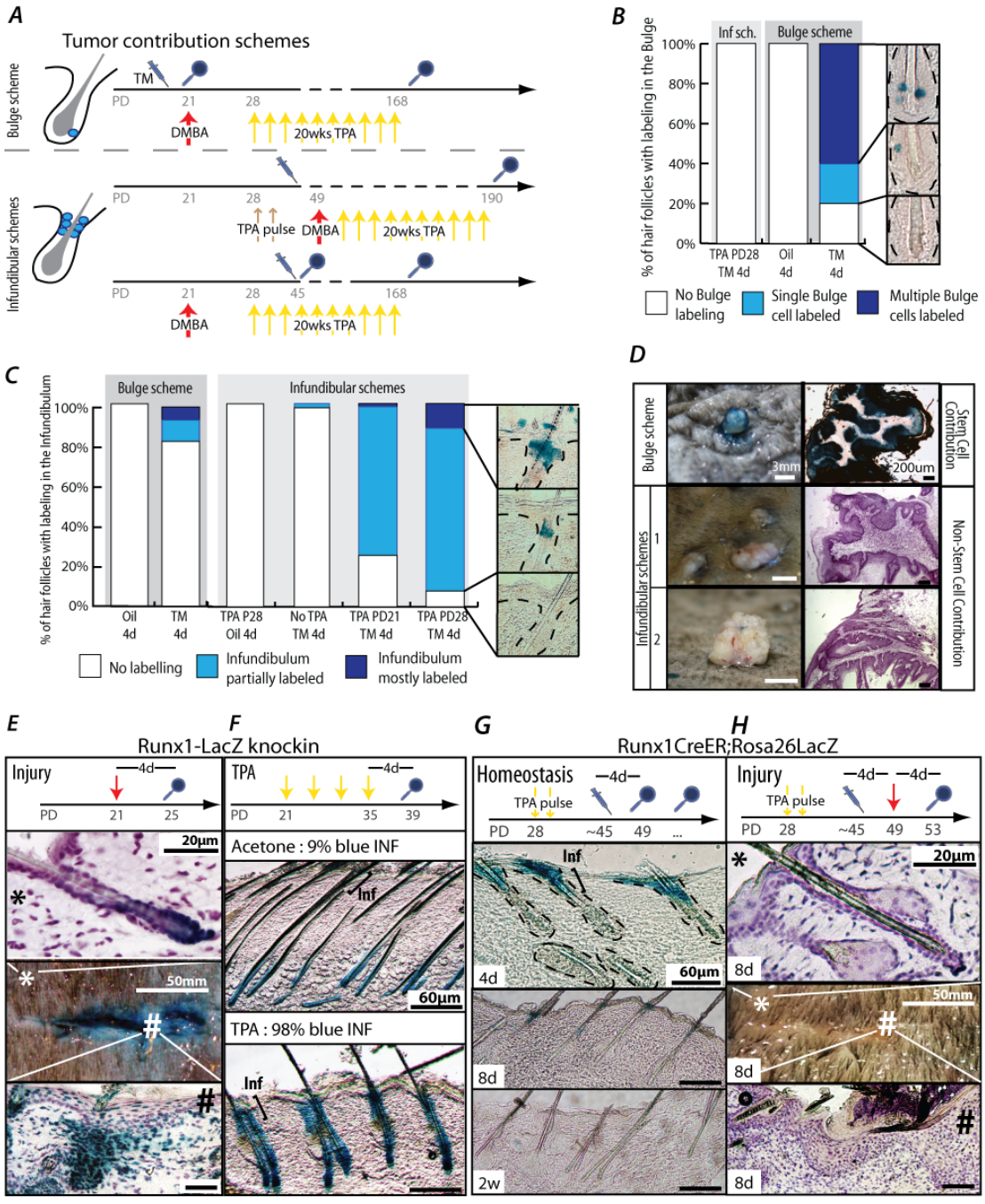

A
Tumor contribution schemes
Bulge scheme
Infundibular schemes
TM
DMBA
20wks TPA
TPA pulse
PD
21
28
45
49
168
190
B
Inf sch.
Bulge scheme
% of hair follicles with labeling in the Bulge
100%
80%
60%
40%
20%
0%
TPA PD28 TM 4d
Oil 4d
TM 4d
No Bulge labeling
Single Bulge cell labeled
Multiple Bulge cells labeled
C
Bulge scheme
Infundibular schemes
% of hair follicles with labeling in the Infundibulum
Oil 4d
TM 4d
TPA P28 Oil 4d
No TPA TM 4d
TPA PD21 TM 4d
TPA PD28 TM 4d
No labelling
Infundibulum partially labeled
Infundibulum mostly labeled
D
Bulge scheme
Infundibular schemes
1
2
3mm
200um
Stem Cell Contribution
Non-Stem Cell Contribution
E
F
Runx1-LacZ knockin
Injury
TPA
4d
PD
21
25
35
39
20µm
50mm
Acetone : 9% blue INF
TPA : 98% blue INF
Inf
60µm
G
H
Runx1CreER;Rosa26LacZ
Homeostasis
TPA pulse
~45
49
53
4d
8d
2w
60µm
20µm
50mm

Table 2-3: Counts of blue/non-blue tumors for lineage tracing of Runx1 expressing bulge or infundibular cells during tumor development

Genotype	***Bulge Scheme***	***Infundibular Scheme 1***	***Infundibular Scheme 2***	***Oil (Bulge Scheme)***
Total # of tumors	89	75	42	71
# of blue tumors	3	0	0	0
% of blue tumors	**3.37%**	**0%**	**0%**	**0%**

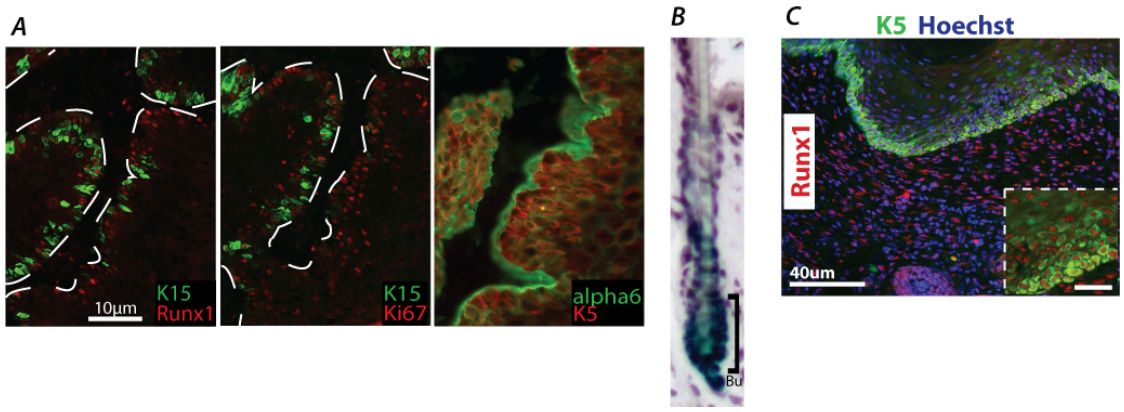


Figure 2-8: Tumors and wounds show similar Runx1 expression patterns

A: Immunofluorescence staining of a fully blue tumor from the bulge lineage tracing. Patterns of K15, K5, alpha6, Ki67 and Runx1 were as previously observed. Dashed line represents the boundary between tumor and stroma where it is not apparent from the staining. B: X-Gal staining of Runx1-LacZ KI mouse showing labeling in the entire bulge (Bu). C: Immunofluorescence staining for Runx1 and K5 at the wound in Runx1-LacZ KI mice.

endogenous Runx1 promoter activity and expression (North *et al.*, 1999) we estimate a conservative 50% of all bulge cells as normally being able to express Runx1 (Figure 2-8B). These should have been labeled by the Runx1-CreER induction and we can thus infer that a significant fraction (~70% if not more), of the skin tumors in this procedure may arise from Runx1+ bulge cells. The incidental labeling in 20% of infundibuli make these cells as potential alternative candidates for the origin of at least some of the tumors obtained in our treatment.

Tissue injury and inflammation has long been recognized as an oncogenic factor. Intriguingly, we found that injury or short-term TPA treatment induced ectopic Runx1 expression in the hair infundibulum and epidermis just adjacent to the HF (Figure 2-7E,F, Figure 2-8C). Micro-injury may be responsible for the 20% infundibular labeling present in our previous experiment. We asked whether the infundibullar cells rendered to express Runx1 upon TPA induction or injury are SCs and may contribute to tissue homeostasis, injury repair, or tumor formation. Injection of TM in Runx1-CreER; ROSA26R mice at the onset of the second telogen (PD38-45) following approximately 2.5 weeks of TPA treatment (Figure 2-7A, Infundibular scheme) resulted in the selective labeling of the majority of infundibular cells in each HF (Figure 2-7B,C,G, Infundibular scheme) but no labeling in the bulge or the interfollicular epidermis (n~1000). Two weeks after TM injection, we sacrificed mice that had been injured by a penetrating cut and examined skin sections by X-gal-staining. We determined that this population contributed neither to HF homeostasis (away from cut) nor to healing of the

epidermal wound (Figure 2-7G,H). Therefore, we conclude that cells ectopically expressing Runx1 in the infundibulum do not acquire tissue SC properties.

We next asked if aberrant expression of Runx1 in these more differentiated HF cells had an oncogenic effect in the experimental conditions of our 2-stage skin carcinogenesis protocol. DMBA treatment 4 days after or 2.5-3 weeks before labeling by TM injection, followed by 20-weeks of TPA to generate tumors did not lead to any fully- or partially-X-Gal+ papilloma (Figure 2-7A,D, Infundibular schemes). These data, together with the fact that we had achieved a much greater labeling efficiency in the infundibulum versus the bulge labeling schemes (Figure 2-7B,C), suggest that the ectopic Runx1-expression in populations of non-SCs do not render these cells capable of forming tumors. Additionally the results indicate that Runx1 is essential for tumor formation only in HFSCs but not sufficient outside this niche, even if those more differentiated short-lived cells had been previously initiated with a mutagenic drug.

2.3.5 Absence of Runx1 leads to tumor regression

The data thus far demonstrated that Runx1 is in the cells of tumor origin, that these cells must be HFSC to generate tumors, and that Runx1 is essential in the initiation stage of tumor formation. Moreover, the loss of Runx1 negative tumors in the absence of a proliferative agent (Figure 2-4H) suggested a possible role for Runx1 in tumor maintenance.

To address this question, we attempted to remove Runx1 from tumors after they

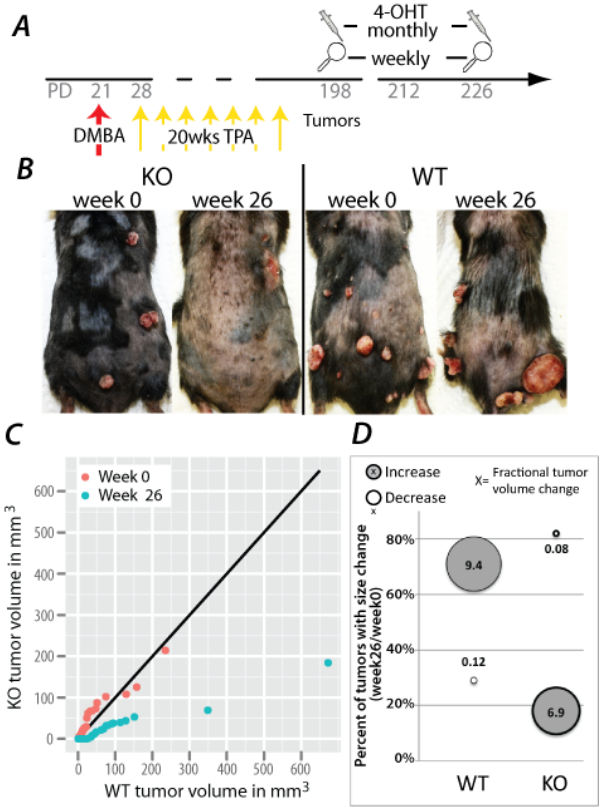


Figure 2-9: Runx1 loss causes tumor regression

A: Scheme of tumor regression study in Runx1 iKO mice. B: Representative photos of Runx1 WT and KO mice at 0 and 26 weeks of 4OHT treatment. C: q-q plot comparing the tumor distributions of WT (x-axis) and KO (y-axis) mice at 0 and 26 weeks. The black line indicates where the two genotypes would have equal distributions. A point to its right denotes more tumors of that size in WT mice and vice versa. D: Quantification of growing or shrinking tumors in WT or KO mice including fractional tumor change (circle size).

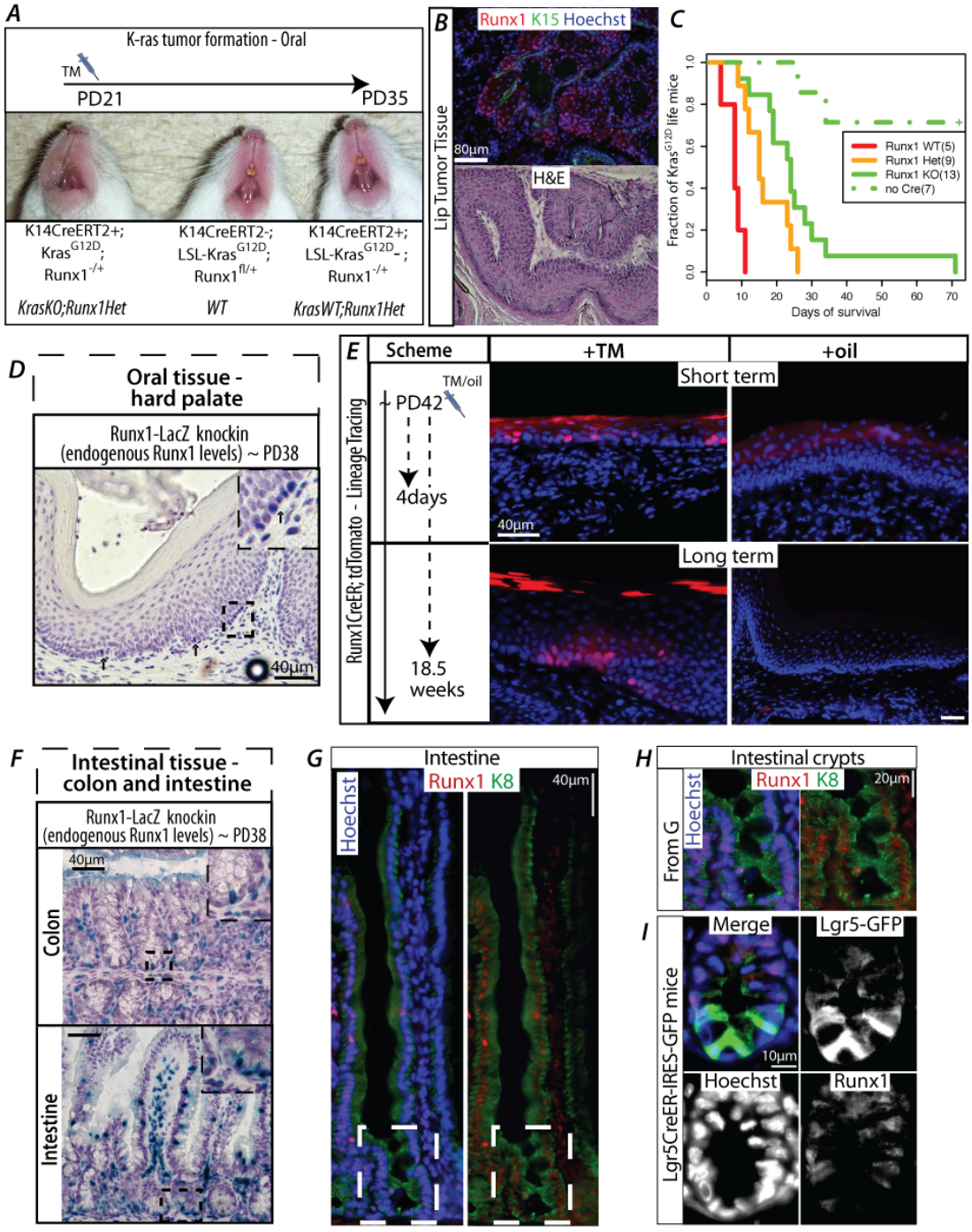
had formed and monitored their number and size over time. To remove Runx1, we used topical applications and monthly injections of 4-OHT in acetone directly into tumors (Figure 2-9A). One month after the first injection Runx1 was lost in 50% of 29 tumors analyzed. We next measured the volume of 72 treated tumors bi-weekly for 26 weeks. We found that ~ 75% of WT tumors (acetone only) increased 940% in size while 80% of 4-OHT injected tumors decreased to 8% of their original size, suggesting that Runx1 loss induced roughly 120-fold shrinkage of the tumor size (Figure 2-9B-D). This is likely an under-estimation, as all remaining tumors in Runx1 KO mice at 26 weeks showed Runx1 expression due to incomplete KO (not shown). In conclusion, the removal of Runx1 from developed epithelial skin tumors drastically reduces the size of these tumors.

2.3.6 Runx1 is important for oral tumor formation in LSL-Kras^{G12D} mice and is expressed by oral epithelial stem cells

Next we aimed to expand our understanding of Runx1 role in stem cells and tumors beyond that of skin SCC and HFSCs. Runx1 appeared highly expressed in oral epithelium human SCC lines and primary tumors (Figure 2-2). DMBA, which induces the initiating mutation in the two-stage carcinogenesis protocols utilized here, has been shown to target the Ras genes with high frequency (Abel *et al.*, 2009) and a considerable fraction of human oral and skin SCCs show mutation in one of the Ras genes (Pierceall *et al.*, 1991; Spencer *et al.*, 1995). Expression of an activated form of K-Ras (KRas^{G12D}) initiates squamous tumor formation in the oral and anal epithelium as well as in the skin (Vitale-Cross *et al.*, 2004; Gilad *et*

Figure 2-10: Runx1 marks oral stem cells and co-localizes with Lgr5 in the intestine

A,B: Representative images and staining of sections from lip tumor tissue from $Kras^{G12D}$ mice 14 days after TM injection. C: Kaplan-Meier survival analysis for K14-CreERT2;$Kras^{G12D}$ mice of all Runx1 genotypes (p-value = 7.06x10-7) and Cre- controls. D: Endogenous expression of Runx1 in oral tissue of Runx1LacZ-KI mice. Black arrows point to some epithelial LacZ+ cells. Inset shows a higher magnification of the marked tissue area. E: Lineage tracing of oral (hard palate epithelium) Runx1 population using Runx1-CreER; tdTomato mice. F: Same as D for colon and intestinal epithelium. G: Runx1 and K8 staining of an intestinal villus and crypt in wild-type mice. H shows the close-up marked by the dashed line in G. I: Runx1 staining in crypts of Lgr5CreER-IRES-GFP mice where GFP expression highlights the intestinal stem cell population.


A
K-ras tumor formation - Oral
TM
PD21
PD35
K14CreERT2+; Kras^{G12D}; Runx1$^{-/+}$
KrasKO;Runx1Het
K14CreERT2-; LSL-Kras^{G12D}; Runx1$^{fl/+}$
WT
K14CreERT2+; LSL-Kras^{G12D}-; Runx1$^{-/+}$
KrasWT;Runx1Het
B
Runx1 K15 Hoechst
Lip Tumor Tissue
80μm
H&E
C
Fraction of Kras^{G12D} life mice
Days of survival
Runx1 WT(5)
Runx1 Het(9)
Runx1 KO(13)
no Cre(7)
D
Oral tissue - hard palate
Runx1-LacZ knockin (endogenous Runx1 levels) ~ PD38
40μm
E
Scheme
+TM
+oil
Runx1CreER; tdTomato - Lineage Tracing
TM/oil
~ PD42
4days
18.5 weeks
Short term
Long term
40μm
F
Intestinal tissue - colon and intestine
Runx1-LacZ knockin (endogenous Runx1 levels) ~ PD38
40μm
Colon
Intestine
G
Intestine
Runx1 K8
Hoechst
40μm
H
Intestinal crypts
From G
Hoechst
Runx1 K8
20μm
I
Lgr5CreER-IRES-GFP mice
Merge
Lgr5-GFP
10μm
Hoechst
Runx1

al., 2010; Lapouge *et al.*, 2011). To connect our findings with genetic pathways of human SCC, confirming that tumors with activated Ras depend on Runx1 for their formation, and to analyze the role of Runx1 in the oral epithelium tumor formation we utilized LSL-KRas^{G12D} mice that can be activated by TM induced Cre activity (Jackson *et al.*, 2001). We generated K14-CreERT2;LSL-Kras^{G12D};Runx1 KO, wild-type, and heterozygous mice and injected TM at PD21 to induce either the Runx1 KO, or K-Ras activation, or both. The Kras^{G12D} mice rapidly developed sublingual and hard palate oral papillomas (Figure 2-10A-C) at low TM doses as previously reported (Gilad *et al.*, 2010). Few K14-CreERT2 negative mice developed tumors as well, indicative of low frequency leaky, non-Cre dependent K-Ras activation. Oral tumor formation severely impaired the mice's ability to take in food, ultimately leading to premature death of these mice, prior to skin tumor formation. Although we observed oral or anal tumors in all Kras^{G12D} mice, tumor formation and subsequent death were significantly delayed in Runx1 KO mice. Moreover a dosage-dependent effect was apparent in mice with a single intact copy of Runx1 (Figure 2-10C). These data extend our results on the importance of Runx1 from skin to oral and anal epithelial tumorigenesis. Moreover, they underscore the function of Runx1 in a genetic context (Kras^{G12D}) highly common in human skin and head and neck (oral) and skin SCC, complementing our 2-step skin carcinogenesis data.

Given Runx1 function in oral tumorigenesis, we examined its expression in the oral epithelium using the Runx1-LacZ KI. At PD38 Runx1 expression is in the

basal layer of the palate epithelium. (Figure 2-10D). This area was previously shown to contain label-retaining cells, which were putative SCs (Bickenbach and Mackenzie, 1984).

Next, we traced the lineage of oral Runx1+ cells using mice containing the inducible Runx1-CreER and the tdTomato reporter (Madisen *et al.*, 2010). Labeling was induced mostly in the basal layer of the hard palate 4-days after TM but not oil injection (Figure 2-10E). At 18.5 weeks after TM injection we detected rare clones consisting of basal layer (un-differentiated) and supra-basal (differentiated) cells (Figure 2-10E). This is the first study showing that indeed a subpopulation of oral basal layer cells differentiate and self-renew for extended periods of time. Oral epithelium SCs express Runx1, making this the third adult tissue SC type with lineage tracing evidence for Runx1 expression, after HFSCs (this work) and HSCs (Samokhvalov *et al.*, 2007).

2.3.7 Runx1 expression in colon and intestine

Because we observed Runx1 expression in several human cancer cell lines of epithelial origin other than skin and oral, we decided to analyze Runx1 endogenous expression in a third tissue: the colon and intestine. We used the Runx1-LacZ KI mice at PD38 and found X-gal signal, and thus endogenous Runx1, in villus cells but more importantly, at the base of every crypt (Figure 2-10F) where intestinal SCs reside (Barker *et al.*, 2008). We confirmed this by staining using a Runx1 antibody in wild-type intestinal tissue (Figure 2-10G,H) and in Lgr5CreER-IRES-GFP mice. Indeed, some Runx1 expressing crypt cells co-localize

with GFP, suggesting they may be SCs (Figure 2-10I). Notably, staining of Runx1 and K8, an epithelial marker, in wild-type mice revealed an overall pattern of Runx1 expression that is similar to HFs: Runx1 is lightly expressed in a few SCs at the base of the crypt (analogous with the hair bulge) and strongly expressed in the activated cells or transit amplifying cells in the upper crypt (analogous with the hair germ). Extremely low labeling efficiency of Runx1-CreER in the intestine and colon with either tdTomato, YFP or ROSA26R reporter prevented lineage tracing of these Runx1+ cells. It is possible that by studying the isoform structure of Runx1 in the intestine and different CreER knockin strategy may overcome this limitation. These data suggest that Runx1 may be expressed in the intestine in a sub-fraction of the LGR5+ stem cells.

2.3.8 Human cancer cells require Runx1 for growth

Thus far we presented evidence that Runx1 is a SC factor for at least two epithelial tissues, and is important for tumor initiation/formation in these tissues. Runx1 is also expressed in several kinds of human epithelial cancer cells (Figure 2-2E) and tumors, as shown by Oncomine data (Figure 2-2A). To test the significance of Runx1 in human epithelial cancers we knocked down its expression in several cancer cell lines and monitored their cell growth and survival.

We used Luciferase shRNA (Luc) as a control and two Runx1 shRNAs. Runx1-1 and Luc expression, but not Runx1-2, can be assessed by detection of GFP (Materials and Methods). Both Runx1 shRNAs (but not Luc) knocked down

Runx1 levels (Runx1KD), albeit to a different degree, as determined by immuno-fluorescence staining with a Runx1-specific antibody (Figure 2-11A,B). Skin and head and neck (oral) SCC cell lines with Runx1KD failed to grow, confirming an essential role for Runx1 in tumor maintenance as seen in mouse papillomas (Figure 2-11C,D and Figure 2-9). Ovarian cell lines show a reduction in growth, whereas colon, breast and prostate cancer cell lines were not affected (Figure 2-11D). Notably, cells react at different degrees to the two Runx1 shRNAs, likely due to variable knockdown efficiencies of the two target sequences. The effects of Runx1KD can be rescued by the over-expression of Runx1 (Figure 2-11C, Wang *et al.*, 2011).

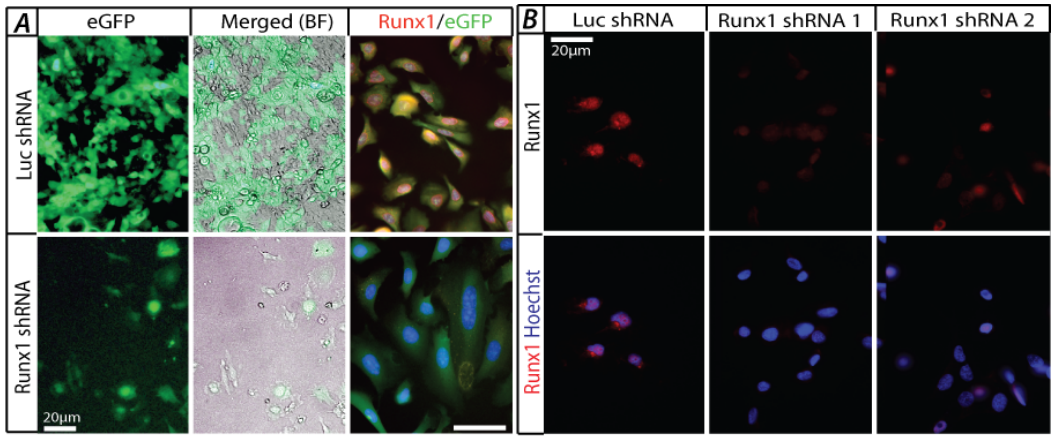
Since all cancer cell lines expressed Runx1 and their tissues of origin also displayed suggestive Runx1 expression we briefly looked into a possible mechanism for resistance to loss of Runx1. In a recent study in breast cancer Runx1 levels were inversely correlated with FOXO1 expression, compensating for Runx1 loss (Wang *et al.*, 2011). However, we could not find a correlation between Runx1KD resistance and FOXO1 in any of the cell lines analyzed here (Figure 2-12A). Overall, these results show that loss of Runx1 is detrimental to growth of three human epithelial cancer cell types: head and neck (oral) SCC, skin SCC, and in part to ovarian cancer cells.

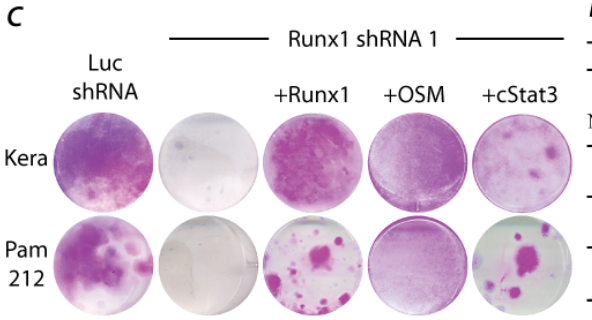
2.3.9 Runx1 acts upstream of Stat3 in cancer cell growth and survival

To understand the mechanism by which Runx1 affects cancer cell growth in mice and humans we considered a possible interaction with Stat3, which appears to

Figure 2-11: Runx1 is important for cancer cell growth by modulating Stat signaling

A: Images of Pam212 cancer cells transfected with control shRNA (Luc) and Runx1-1 shRNA reveal severe growth impairment in the latter (left and middle). Single colony detected with low levels of Runx1 (right) demonstrate the specificity of our antibody and the down-regulation of Runx1. B: Runx1 levels in MDA-231 cells with Runx1 shRNA 1 and 2. C: Representative images of cell plates stained with Rhodanile Blue of Keratinocytes (Kera) and Pam212 cancer cells transfected with control or Runx1 shRNA. Note severe impairment in growth in the Runx1 KD cells. Cell growth was rescued with over-expression of Runx1, addition of OSM, and cStat3. D: Summary of KD experiments in human cell lines and rescue with OSM where applicable. E: Immunostaining of pStat3T705 levels in tumors with or without Runx1 expression. F: Western blots of iKO keratinocytes treated either with ethanol (WT) or 4-OHT for 4 or 5 days. G: QRT-PCR of Runx1 RNA levels in the cells from D (average ± SD). H:Western blot of iKO (4d) cells treated with OSM for 1 day. I: Volcano plot of Jak/Stat pathway qPCR array. J: Scheme for SOCS ChIP primers color coded by the p-value determined for Runx1 binding (grey >0.07, yellow >0.05, green <0.05). K: ChIP of WT and iKO cells (4d) using Runx1 antibody for SOCS family members (average ± SD). P-values for pair-wise comparisons are indicated. Significant p-values are highlighted in green. L: Quantification of GFP+ cells in Keratinocytes and Pam212 transfected with control or Runx1 shRNA and pLKO1-puro (control) and 2 shRNAs each for SOCS3 and SOCS4 (average ± SD). Note lack of GFP+ cells when Runx1 is KD and their rescue by concomitant KD of Runx1 and either SOCS3 or SOCS4.

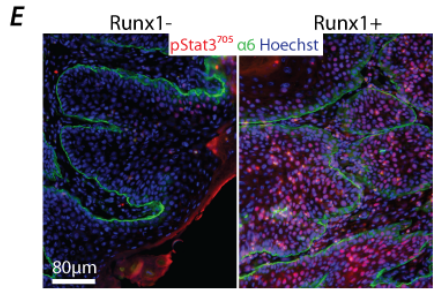


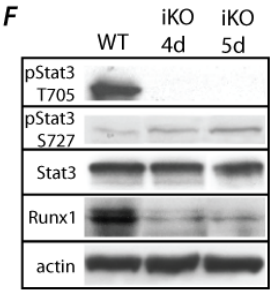


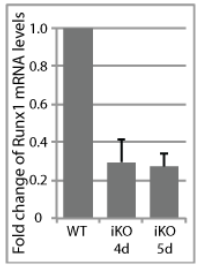
D

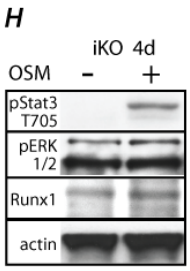
Organ	Cell Line	shRNA 1	shRNA 2	+ hOSM
Skin SCC	SCC13	○	○	●
Head and Neck SCC	SCC66	●	○	◉
	SCC74	○	○	●
	SCC125	●	○	●
Ovarian Cancer	SKOV3	●	◌	◉
	OVCAR3	●	◌	◉
Colon Cancer	HCT116	●	●	
	CCSC	●	●	
Prostate Cancer	DU145	●	●	
	PC3	●	●	
Breast Cancer	MDA231	●	●	
	MCF7	●	●	

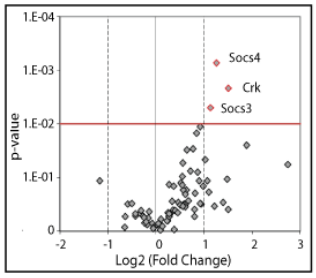
● 100% Growth
◉ 80% Growth
○ No Growth
◌ Little Growth

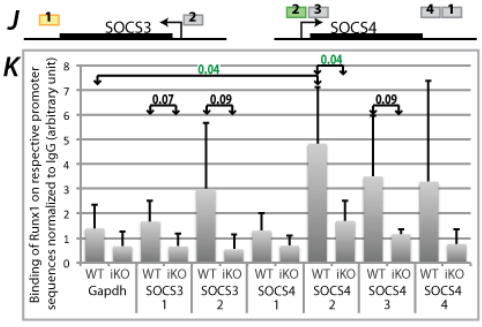


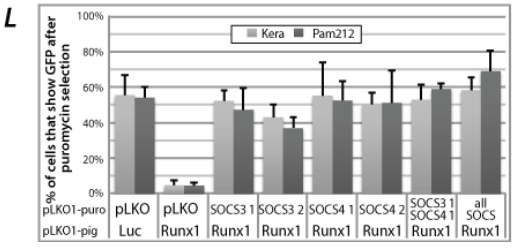












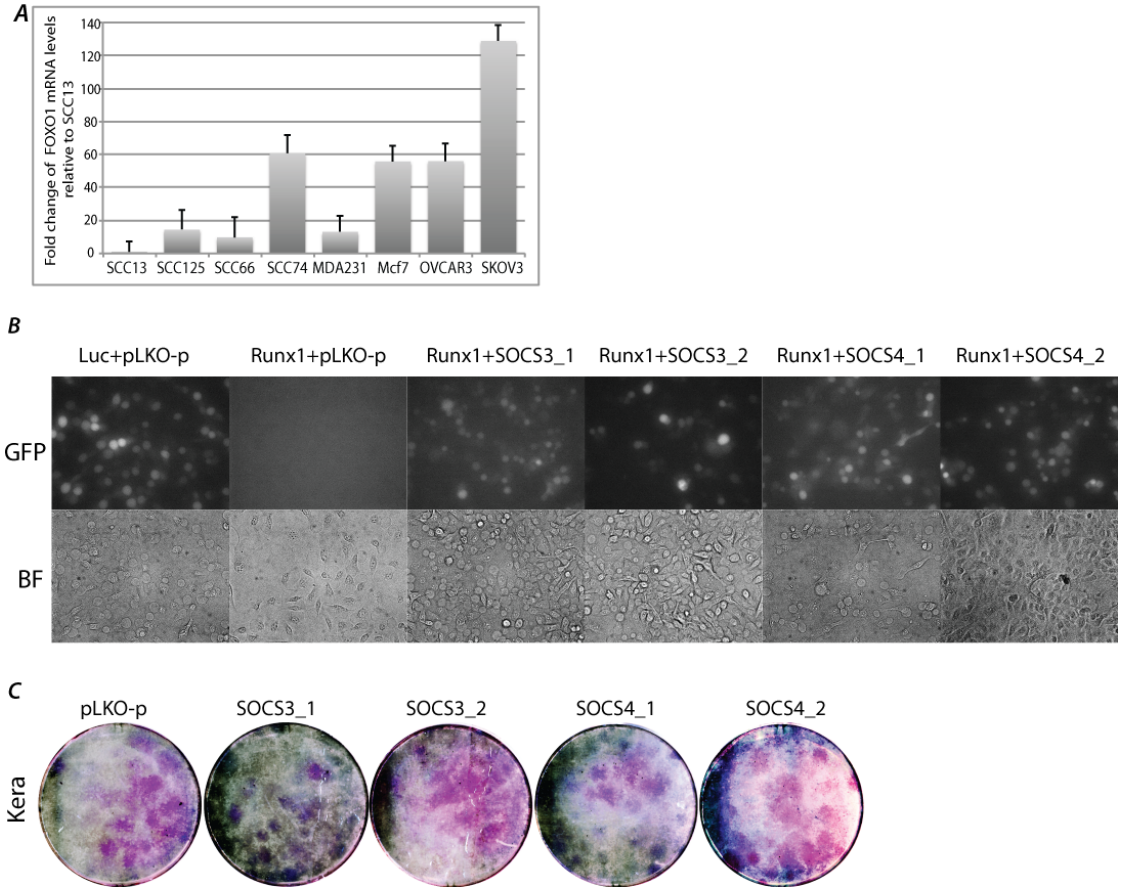


Figure 2-12: FOXO1 and Runx1 levels do not correlate

A: FOXO1 expression in selected human cancer cell lines normalized to SCC13. Error bars stem from replicate experiments. B: Sample bright field and GFP images from SOCS rescue assay in keratinocytes. Note absence of GFP signal in Runx1 shRNA and pLKO-puro backbone co-transfection and rescue of GFP positive cells in Runx1 and SOCS shRNA co-transfection similar to levels of control cells (Luc shRNA + pLKO-puro backbone). GFP negative cells contain pLKO-puro backbone or SOCS shRNA respectively. C: Rhodanile blue staining of pLKO1-p or SOCS shRNA transfected Keratinocytes (Kera) after 5 days of puromycin selection. There is no observable difference between confluency of the different cells.

play a similar role as Runx1 in regulating the hair cycle, and in tumor formation (Sano *et al.*, 1999, 2000; Chan *et al.*, 2004a, 2004b; Osorio *et al.*, 2008; Kim *et al.*, 2009). The activated form of Stat3 - pStat3 T705 - is up-regulated in skin upon TPA treatment in WT mice (Chan *et al.*, 2004a). This up-regulation appeared more variable and overall reduced in the Runx1 KO skin (Hoi et al 2010), and is lost in tumors that form after loss of Runx1 during the tumor promotion stage (Figure 2-11E). Given the essential role of Stat3 in proliferation and survival of human and mouse SCC (Bito *et al.*, 2011; Rho *et al.*, 2011), this may explain why the Runx1 deficient tumors obtained here in the promotion scheme were short-lived and could not be maintained in the absence of TPA (Figure 2-4H). Given the complex role of the microenvironment in regulating Stat3 activation, the manner in which Runx1 interacts with the Stat3 signaling pathway, and the functional relevance of this interaction to Runx1 mechanism of action are unclear. For example, Stat3 down-regulation in the Runx1 KO skin treated with TPA could be an indirect effect to reduced inflammation potentially driven by a paracrine function downstream of Runx1, rather than a direct interaction of Runx1 with the Stat3 pathway itself.

To analyze whether Runx1 directly regulates Stat3 activation, we utilized cells removed from the complex skin and tumor environment and created inducible KO K14-CreER;Runx1$^{fl/fl}$ keratinocytes (iKO). After 4 days of 4-OHT treatment, Runx1 mRNA and protein level were significantly reduced as shown by Western blot and by QRT-PCR (Figure 2-11F,G). Moreover, Western blots for Stat3, pStat3

T705, and pStat3 S727 revealed a selective and striking de-phosphorylation of T705 upon Runx1 loss (Figure 2-11F). In this form Stat3 is inactive because it cannot dimerize and translocate to the nucleus (Ihle, 1996).

To analyze if the disruption of the Jak/Stat pathway is responsible for the growth defect of normal and cancer cell lines with Runx1 deficiency, we attempted to restore Stat3 activity in Runx1KD cells. To this end, keratinocytes and Pam212 cells were either co-transfected with Runx1-1 shRNA and cStat3, the constitutively active form of Stat3 (Bromberg *et al.*, 1999) or treated with Oncostatin M (OSM), a small signaling peptide that activates Stat3 signaling (Ishihara and Hirano, 2002). In both cases, the Runx1 deficient cell lines were then able to form large colonies (Figure 2-11D), and showed increased levels of T705 pStat3, but not Runx1, relative to untreated cells (Figure 2-11H).

OSM activates Jak and Tyk kinases through its receptors that in return activate the Stat3 and the MAP kinase pathway (Erk1/2) (Ishihara and Hirano, 2002). To see if the effect of OSM can be attributed to Stat3 or whether Erk phosphorylation plays a role in the rescue of growth defects in our cells, we examined P-Erk1/2 level by Western blotting. We fail to see an increased phosphorylation of Erk1/2 in keratinocytes (Figure 2-11H), which suggests that it is the Stat3 activation that rescued the Runx1 deficient growth defect, further supporting the similar effect obtained with the cStat3 (Figure 2-11C).

Finally, we employed the human cancer cell lines that previously showed dependence on Runx1 levels, and found that OSM rescues their growth to

different degrees (Figure 2-11D), with the most striking effects on skin and head and neck SCC lines and a milder effect on ovarian cancers. Hence, in normal keratinocytes as well as in three epithelial cancer cell types Runx1 loss appears to impair growth by directly preventing Stat3 activation.

2.3.10 Runx1 stimulates the Jak/Stat pathway by repression of SOCS3 and SOCS4

Next, we attempted to understand the mechanism by which Runx1 may stimulate Stat3 signaling. Since our Western blots suggested stable Stat3 protein levels upon Runx1 loss, it seemed likely that some members of the Jak/Stat pathway involved in T705 phosphorylation might be Runx1 transcriptional targets. To find such putative targets, we surveyed the status of the Jak/Stat pathway using SA-Biosciences RT2 Profiler PCR array and compared mRNA levels in Runx1-CreER;Runx1$^{fl/fl}$ and Runx1$^{fl/fl}$ keratinocytes after 4 days of 4-OHT treatment. A volcano plot showing the fold changes between WT and KO versus the associate p-value for each tested gene revealed 3 genes significantly up-regulated in the iKO cells: Crk, SOCS3 and SOCS4 (Figure 2-11I, raw data in Appendix 1,Appendix 2). Crk is a proto-oncogene thought to stimulate proliferation (Tanaka, 2009), and as such its up-regulation cannot explain proliferation defects but is likely an attempt to compensate for Runx1 loss. On the other hand, SOCS3 and SOCS4 suppress Stat3 T705 phosphorylation via interference with Jak2 and EGFR kinase activity respectively (Larsen and Ropke, 2002; Segatto *et al.*, 2011).

To test if Runx1 directly binds to the SOCS3 and SOCS4 promoters to represses their transcription, we employed two independent K14-CreERT2;Runx1$^{fl/fl}$

keratinocyte lines and performed chromatin immunoprecipitation (ChIP) using Runx1 specific antibodies in KO and WT ($Runx1^{fl/fl}$) cells. We followed with QRT-PCR using primers encompassing SOCS3 and SOCS4 genomic regions containing Runx1 binding sites (Figure 2-11J). Our data showed enrichment from one primer pair in each of the SOCS4 and SOCS3 gene loci (p=0.04 and 0.07 respectively). We used a region of the Gapdh promoter free of conserved Runx1-binding sites as a negative control (Figure 2-11K). Binding of Runx1 on the SOCS3 gene promoter, although here showed a p value slightly lower than the conventional 0.05, was also reported by ChIP Seq data on a blood cell line (Wilson *et al.*, 2010). Thus we suggest that Runx1 represses SOCS3 and SOCS4 expression, likely via direct binding to their promoters, and ultimately up-regulates Stat3 activity by enhancing phosphorylation at T705.

If Runx1 loss induced defects were mediated through SOCS3 and/or SOCS4, as indicated by the QRT-PCR and ChIP experiments described above, simultaneous knockdown of SOCS3 or 4 and Runx1 should rescue the cell growth defect. This could be measured by the presence of GFP+ cells, which indicate Runx1 deficiency (Materials and methods). As expected, cells transfected with control Luc shRNA but not those transfected with Runx1-1 shRNA and the empty plasmid backbone (pLKO1) did not show a population of GFP+ keratinocytes or Pam212 cells after puromycin selection. In contrast, addition of either SOCS shRNA instead of the empty plasmid backbone rescued the GFP+ (Runx1 deficient) population, in which cells with high SOCS were able to grow without Runx1

(Figure 2-11L, Figure 2-12B). Knockdown of SOCS3 or 4 alone did not affect cell growth as expected from low or absent SOC3/4 expression in WT cells and as shown in Figure 2-12C and through equal amounts of GFP- cells in Luc+backbone or Runx1+SOCS shRNA transfected cells. These data together suggested that Runx1 is a direct transcriptional repressor of SOCS4 and potentially SOCS3, which results in activation of Stat3 and thus allows growth of normal and cancer cells (Figure 2-13). In conclusion, Runx1 repression is accompanied by activation of SOCS genes, inhibition of Stat3 activity, and growth impairment of some epithelial cancer cells.

2.4 Discussion

Here we uncover Runx1, as a broader epithelial tissue stem cell and tumor factor than previously recognized. We not only confirmed its expression in adult mouse HFSCs by direct lineage tracing, but also we implicate Runx1 in previously uncharacterized oral epithelium stem cells. We showed that mouse oral epithelium tumors, like their skin counterparts (Hoi *et al.*, 2010 and this work), also appear dependent on Runx1 for their growth. We find that increased proliferation due to ectopic injury-induced expression of Runx1 in skin non-stem cells does not render those cells tumorigenic, while Runx1-expressing stem cells are at the tumor origin. We implicate Runx1 as a key molecule upstream of a central epithelial cancer pathway, Stat3 (Li *et al.*, 2011), which Runx1 activates by direct transcriptional repression of SOC3/4. We documented for the first time the expression of Runx1 in primary skin and oral human epithelium SCC and

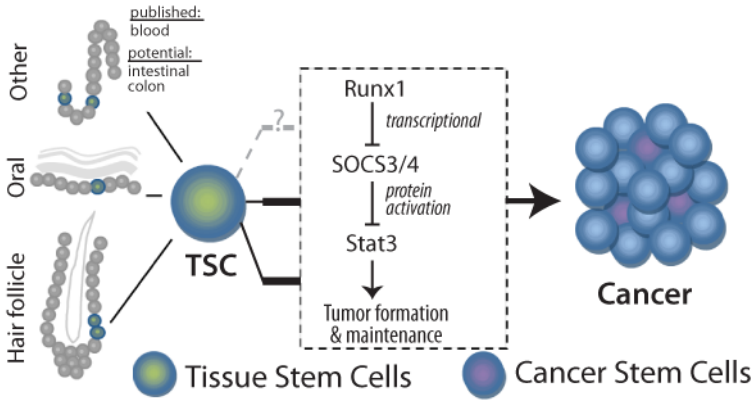


Figure 2-13: Model

Model of Runx1 in tissue stem cells and its role in tumor formation regulation through Jak/Stat family members. Thick black bars highlight conclusions from this paper. Grey dashed line indicates a potential connection.

suggest that several more human epithelial cancers may express Runx1. We demonstrate that 3 different human cancer cell types (skin SCC, oral SCC, and ovarian) depend upon Runx1 for their normal growth, We find Runx1 important for both skin tumor initiation and maintenance, while we previously showed that Runx1 was not essential for adult tissue maintenance (Osorio *et al.*, 2008). This differential requirement for Runx1 places it as a promising target for skin, and potential other epithelial cancers.

Runx1 has been intensively studied for the past 4 decades as a cancer gene frequently mutated in acute myeloid leukemia, which when studied in mice turned out to work as a master regulator of hematopoiesis (Mangan and Speck, 2011). Taking the opposite direction- from mouse to human- we first demonstrated that Runx1 is important for the timely emergence and maturation of HFSCs in embryogenesis (Osorio *et al.*, 2011) and the activation and proliferation of adult HFSCs (Osorio *et al.*, 2008). Moreover, Runx1 appears essential in mice for skin squamous epithelial tumor formation (Hoi *et al.*, 2010). Importantly, with this work we now take Runx1 back to humans where we analyzed the role of Runx1 in several epithelial cancers.

Runx1 surfaced among the top 10% highly expressed genes in 1/3 of all 138 applicable microarrays studies of human cancers found by us in the Oncomine database. These included the skin SCC and BCC, oral SCC, breast, esophageal, lung, colon, and pancreatic cancer among others. We detected Runx1 expression by immuno-fluorescence staining of several kinds of skin human primary tumors as

well as in oral epithelium SCCs. Moreover, we found Runx1 over-expressed in a whole panel of established human cancer cell lines. Importantly, Runx1 appears to be required for growth of skin and head and neck (oral) SCC, and to some extent ovary cancer cells, suggesting Runx1 as a novel putative target for treating these cancers. In contrast, although Runx1 was expressed at high levels it was dispensable for growth of breast, prostate and colon cancers. In light of a recent study of breast cancer, these cells may heave learned to live without Runx1 by up-regulating compensatory pathways (Wang *et al.*, 2011). However, it appears that FOXO1, shown to compensate for Runx1 loss in MCF10A-5E breast cancer cells, is not up-regulated in the cells we analyzed here. This information will be important for future treatment of some cancer types in which a combination drug therapy against Runx1 and compensatory pathways may be required.

An attractive clue positioning Runx1 as a promising therapeutic target is our demonstration that loss of Runx1 shrinks tumors in mice and impairs human and mouse tumor cell growth *in vitro*. This potential adds an important angle to the role of Runx1 in tumor initiation, which could be exploited for tumor treatment and prevention respectively. Runx1 loss is not essential for normal skin and oral epithelium homeostasis, and deletion of Runx1 in adult mice via a ubiquitous β-Actin-CreER allele or Mx1-Cre had mild effects on the mouse physiology (Appleford and Woollard, 2009; Hoi et al., 2010; and our unpublished data). The tolerance of the tissue to Runx1 loss in HFSCs is probably due to the unique protective niche environment, which is disrupted in cell culture and in tumors.

Disrupting Runx1 binding to the DNA via specific drugs is already considered as a potentially beneficial approach to treating leukemia (Mangan and Speck, 2011). A more targeted approach for local treatment of accessible skin or head and neck tumors would be even less invasive.

How is Runx1 working in regulating tumor formation and maintenance? One possibility is that Runx1 is expressed in the cell of tumor origin and remains expressed in some of the progeny cells subsequently generated. At least in the skin, HFSCs are known to be at the origin of some SCC (Youssef *et al.*, 2010; Lapouge *et al.*, 2011; White *et al.*, 2011). Previously, we reported Runx1 expressed in a fraction of cells in bulge and hair germ cells, the regions of the HF where SCs are known to reside. However, it is unclear whether all cells in these regions are SCs. Even though loss of Runx1 impairs proliferation of HFSCs it was possible that Runx1 modulates the SC environment from neighboring cells. For example, Runx1 is able to regulate production of several secreted Wnt molecules during embryogenesis (Osorio *et al.*, 2011). Our lineage tracing work directly demonstrates here for the first time that adult Runx1-expressing cells are long-lived, self-renewing and differentiating HFSCs, which are at the tumor origin. Moreover, Runx1 also appears to be a marker for oral epithelium SCs. Our data suggests some overlap of Runx1 and LGR5 expression, and LGR5 is a well-recognized intestinal SC marker. Further substantiation by lineage tracing is needed for the intestine since the existing Runx1-CreER is extremely inefficient in this organ.

Using skin as a model system we demonstrate that Runx1-expressing HFSC are at the origin of tumors, and these tumors were likely monoclonal. It is recognized that epidermal cells may form skin SCC (Lapouge *et al.*, 2011), and even though Runx1 is not expressed in the epidermis in normal conditions it appears that in our experimental conditions all mouse tumors expressed Runx1. We therefore hypothesized that a population of epidermal cells rendered by inflammation or injury caused by the carcinogenic treatment to ectopically express Runx1 might also generate tumors. However, we found this population located in the infundibulum of the HF to be short-lived progenitors that did not contribute to long-term homeostasis and injury repair. Moreover, this population was unable to generate tumors in our experimental conditions, even in the context of an initiating mutation induced by drug-treatment. This underscores the crucial importance of the intrinsic epigenetic context or cellular make-up to initiation of carcinogenesis and adds to the theory that SCs, and not differentiated cells, are the most likely initiators of cancer (Magee *et al.*, 2012).

Runx1 is known to regulate cell survival, proliferation, differentiation and cell cycle progression (Blyth *et al.*, 2005; Mikhail *et al.*, 2006; Friedman, 2009). Moreover, Runx1 directly binds and represses the promoter of cyclin-dependent kinase inhibitor p21 in skin epithelial cells (Lee, J. and Tumbar, T. submitted; Hoi *et al.*, 2010). Thus, it seemed surprising that Runx1 was essential for tumor initiation but dispensable for tumor promotion, in which over 50% of the tumors in Runx1 KO mice actually lacked Runx1. However, these tumors were lost after

we ceased the TPA treatment, demonstrating that during the proliferative/maintenance stage other strongly proliferative agents can compensate for Runx1 loss. These short-lived Runx1-deficient tumors displayed down-regulation of activated Stat3 and a striking absence of CD34+ cells, which were previously deemed cancer SCs, important for tumor growth and maintenance (Malanchi *et al.*, 2008). These data together, may suggest that Runx1 is expressed in the HFSCs that generate the tumors, and is continuously required in these tumors for the maintenance of their long-lived tumor SC population and for continuous tumor growth. Moreover, Runx1 role in this process is likely mediated via stimulation of Stat3 signaling.

Phosphorylation of Stat3 at T705 and S727 are necessary for full activity of Stat3 (Yokogami *et al.*, 2000). However, phosphorylation at S727 alone prevents activation at T705 and thus locks Stat3 in its inactive state in the cytoplasm (Gartsbein *et al.*, 2006). Upon Runx1 loss we observed a loss of T705 phosphorylation while S727 maintained its levels, indicating a double-safe mechanism to keep Stat3 inactive. Active Stat3 is associated with psoriasis in humans and mice (Sano *et al.*, 2005) and its increased activity has been repeatedly associated with skin SCC formation (Chan *et al.*, 2004a, 2004b, 2008; Aziz *et al.*, 2007). Like Runx1, Stat3 is also considered essential for tumor initiation and maintenance by directly targeting the HFSCs in mouse models (Kim *et al.*, 2009). Activated Stat3 is found in the vast majority of human oral cancers as well as in breast, lung, ovarian, pancreatic and prostate cancers and myelomas,

leukemias and lymphomas (Yu and Jove, 2004). High levels of activated Stat3 have also been reported in oral SCC cell lines, and proliferation of these SCC cells could be further increased with OSM (Douglas *et al.*, 2004), a potent activator of the Stat3 pathway. This connects well to our finding that addition of OSM rescues the proliferation defect of skin and oral SCC and ovarian cancer cells upon Runx1 loss, through activation of Stat3.

Here, we expose a novel regulatory pathway for Stat3 activity. We show that Runx1 regulates Stat3 activation by transcriptional repression of Stat3 inhibitors SOCS3 and SOCS4. Conversely, loss of Runx1 results in up-regulation of SOCS3 and SOCS4 and Stat3 inactivation, which is detrimental to cancer cell growth. Importantly, we show that this mechanism is conserved in humans where it affects skin SCC, and head and neck (oral) SCC, and ovarian cancer. The connection between Runx1 and Stat3 is a putative axis in regulation of cell proliferation and tumorigenesis, suggesting a potential benefit for investigating Runx1 status in Stat3-dependent cancers. If Runx1 is upstream of Stat3 and potentially other major cancer pathway, such as Wnt signaling as we previously showed (Osorio *et al.*, 2011), it may be more beneficial to target the master regulator, Runx1, rather than its individual target signaling pathways.

2.5 Materials and Methods

2.5.1 Samples

Human samples were obtained de-identified and stained as described in the supplement. K14-CreERT2, Runx1$^{fl/fl}$, Runx1-CreER, LSL-Kras^{G12D}, Lgr5CreER-IRES-GFP, ROSA26R, and tdTomato mice were crossed to obtain the various experimental mice. CreER activation by Tamoxifen (Sigma) and 2-step carcinogenesis DMBA/TPA protocol were previously described (Hoi *et al.*, 2010 and see below).

2.5.2 Mice

The Cornell University IACUC approved all the mouse work.

To create Runx1 KO mice, hemizygous K14CreER-T2 mice were mated with Runx1$^{fl/fl}$ mice; F_1 K14CreER-T2;Runx1$^{fl/+}$ offspring were subsequently bred again with Runx1$^{fl/fl}$ mice, generating K14CreER-T2;Runx1$^{fl/fl}$ mice at a 25% Mendelian frequency. This progeny was crossed once more with Runx1$^{fl/fl}$ mice to generate K14CreER-T2;Runx1$^{fl/fl}$ and Runx1$^{fl/fl}$ mice at a 1:1 ratio. Mice were crossed into the FVB background by mating K14CreER-T2;Runx1$^{fl/+}$ to FVB WT mice for 4 generations. K14CreER-T2;Runx1$^{fl/+}$ progeny were then inter-mated with Runx1$^{fl/+}$ progeny to obtain experimental K14CreER-T2;Runx1$^{fl/fl}$ in the FVB background.

To label Runx1 expressing cells, hemizygous Runx1-CreER mice were mated to homozygous Rosa26R-LacZ (BL6) mice. All mice were housed in cages with same sex littermates post-weaning at PD17 or PD21.

For Runx1 removal during tumor formation K14CreER-T2;Runx1$^{fl/fl}$ mice were injected once with 115 µg TM per g of body weight of TM dissolved in corn oil. Runx1$^{fl/fl}$ +TM and K14CreER-T2;Runx1$^{fl/fl}$ +oil served as control mice. For tumor regression 4-OHT was dissolved in acetone to a concentration of 1mg/100ul and applied relative to tumor size (minimum 10ul; maximum 200ul).

In the genetic SCC model K14CreER-T2;LSL-Kras^{G12D};Runx1$^{fl/fl,\ fl/+,\ +/+}$ mice were injected at PD21 with either 225µg/g, 115µg/g, 55µg/g, or 27µg/g TM.

Runx1-CreER mice with ROSA26R or tdTomato (JAX 7914) reporters injected with TM or with oil were used to trace the lineage of Runx1 expressing cells. For long-term HF lineage contribution, mice were injected with 75µg/g at PD17. To trace the origin of skin tumors, mice were injected with 225µg/g once daily on two consecutive days at either PD17/18 (bulge labeling) or 2nd telogen onset (infundibular labeling).

Lgr5CreER-IRES-GFP were not injected with TM and sacrificed at 7 weeks of age. Samples were fixed in 2% PFA for 10 minutes, washed in PBS and embedded in OCT for immunofluorescence staining.

2.5.3 DMBA/TPA treatments

We applied 32µg 9,10-Dimethyl- 1,2-benzanthracene (DMBA) dissolved in 200µl acetone once to the shaved back skin of mice to induce point mutations in the DNA. For the following 20 weeks, the phorbol ester 12-O-tetradecanoylphorbol-13-acetate (TPA) was used twice weekly at a dose of 12.4 µg TPA/mouse dissolved in 200uL acetone to induce robust proliferation of the skin epithelial cells and thus promote the mutation. The appearance of tumor on the back skin and their number in each mouse were counted once a week during the treatment period.

To assess tumor maintenance aged mice between 7-14 months were monitored bi-weekly for tumor size. Tumors and skin sections were embedded in OCT and analyzed.

2.5.4 Skin injury

Wounds were performed either side of the midline using a dissection scalpel. Lineage contribution to wound healing was assessed 4 days, 2 weeks or 5 weeks after injury.

2.5.5 Histology and immunofluorescence

Staining of skin tissue for immunofluorescence and for Hematoxylin and Eosin (H&E) were as described previously (Tumbar, 2006). MOM Basic Kit (Vector Laboratories) was used for mouse antibodies. Nuclei were labeled by Hoechst.

Antibodies for immunostaining: (1) rat [α6-integrin (1:100), CD34 (1:150) (BD Biosciences), E-cadherin (1:500, Sigma) and BrdU (1:300, #ab6326)]; (2) guinea pig [K5 (1:5000, E. Fuchs, Rockefeller)], (3) rabbit β-gal (1:2000, Cappel), Runx1 (1:8000, Dr. T. Jessell, Columbia), Runx1 (1:500, #ab68939), Runx1 (1:300, SDIX generated No25), Ki67 (1:200; Novocastra), and Keratin 14 (1:1000, Covance),]; (4) mouse [AE13 (1:50, Immunoquest), AE15 (1:10, T.T. Sun, NYU), K10 (1:400, ab9025), Ki67 (1:100, Vector labs), K14 (1:100, #ab7800)]; (5) chicken [GFP (1:100, #ab13976), K8 (1:250, #ab14053), K15 (1:5000, #ab77832)]. Secondary antibodies conjugated to either TxR, FitC or Cy5 were used to visualize the signal.

Staining intensity of individual nuclei was quantified with Autoquant. Cells were marked and the integrated voxel intensity was recorded and averaged over all cells counted. Staining intensities of Ki67, caspase3, CD34 and Runx1 were quantified per tumor section with ImageJ.

2.5.6 X-gal staining and whole-mount

For 5-bromo-4-chloro-3-indoxyl-beta-D- galactopyranoside (X-Gal) staining of thin sections, 10 μm skin sections were fixed for 1 minute in 0.1% glutaraldehyde and washed in PBS. Incubation in X-gal solution was at 37°C for 12-16 hours. Thick skin sections, 60-90 μm, were stained as described previously (Zhang *et al.*, 2009). For whole-mount staining animals were sacrificed and the dorsal skin was removed, cleaned of muscle and adipose tissue and fixed 30 minutes in Periodate-Lysine-Paraformaldehyde (McLean and Nakane, 1974), washed with PBS, and stained 48hrs at 4°C with X-gal buffer (Kopan *et al.*, 2002).

2.5.7 Cell Culture

Keratinocyte cell lines were established and grown as described previously (Tumbar, 2006). To induce recombination in iKO cells 2μM/ml 4-OHT dissolved in Ethanol were added. The mouse skin SCC cell line Pam212 was grown in EMEM (Lonza) + 10% FBS. The human skin SCC cell line SCC13 was grown as described in (Rheinwald and Beckett, 1980, 1981). Human oral SCC cell lines SCC66, SCC74 and SCC125 are grown in L-glutamine, non-essential amino acids, and gentamycin supplemented MEM with 10% FBS. Instructions to grow the human colon cancer stem cell line CCSC can be obtained from Dr. Xiling Shen (Cornell). All other human cancer cell lines are grown in media according to instructions from the ATCC: HCT116 , SKOV3,OVCAR3,PC-3, DU-145, MCF-7, MDA-MB-231. All cells are grown at 37C and 7.5% CO_2.

2.5.8 shRNA sequences

Runx1-1: 5'-caagaccctgcccatcgct-3', Runx1-2 [TRCN0000013660]:

5'-cctcgaagacatcggcagaaa-3', Luciferase (Luc): 5'-cttacgctgagtacttcga-3', SOCS3-1 [TRCN0000231176]: 5'-tcttcacgttgagcgtcaaga-3', SOCS3-2 [TRCN0000067468]: 5'-gctaggagactcgccttaaat-3', SOCS4-1 [TRCN0000351105]: 5'-tatttgcagaacggttatttg-3', SOCS4-2 [TRCN0000340013]: 5'-attcagcgcaggaagattatt-3'.

2.5.9 shRNA and virus production

shRNA was either in the pLKO1-puro backbone (Runx1-2; SOCS3-1,2 and SOCS4-1,2) or in a modified version containing a "Puro-IRES-GFP" cassette (pLKO1-pig) (Runx1-1, Luc) (Ang *et al.*, 2011). pLKO1-pig and shRNA for luciferase were a gift from Dr. Michael Rendl. For sequence information see the Supplementary Materials.

To produce shRNA lentiviruses, the construct was co-transfected with packaging (psPAX2, Addgene plasmid 12260) and envelope (pmd2.g, Addgene plasmid 12259) plasmids into HEK293T cells (100mm plate 50% confluent) using Mirrus Transit-LT1 transfection reagent, according to the manufacturer's instructions. Approximately 16 hours after transfection, fresh media was added to the 293T cells and they were cultured for 72 hours. Media was collected and refreshed at 48 hours, and then collected again at 72 hours and concentrated using a 30kd MWCO centrifugal spin column. Concentrated media containing lentiviral particles was aliquoted and stored frozen at -80C (Chuang *et al.*, 2010).

2.5.10 Knockdown

Mouse cells were grown to 50-60% confluency and transfected with Opti-MEM (GIBCO), TransFectin(Biorad) and 1ug of shRNA. Human cells were grown to 60% confluency and polybrene (8ul/ml) was added prior to transduction. Virus based on the pLKO1-pig backbone containing luciferase or Runx1 shRNA was added at about $1x10^6$ particles per dish. Selection with puromycin started at 90%

confluency and was continued for 1-2 weeks, depending on the growth rate of the cell line until cells are fixed and analyzed.

2.5.11 Rescue assays

All knockdown and rescue assays were duplicated or triplicated. Mouse keratinocytes or Pam212 were co-transfected with 1μg DNA of each shRNA (Runx1-1 or Luc in pLKO1-pig) and cStat3 (Addgene plasmid 8722; (Bromberg *et al.*, 1999). At 90% confluency we select for shRNA containing with puromycin. Oncostatin M (OSM) (Cell Signaling human: #5367; mouse: #5371) was used at 10ng/ml media at the time of puromycin selection. pLKO1-puro only or mouse specific SOCS3 or 4 shRNA was co-transfected with pLKO1-pig Runx1-1 and Luc shRNA. All transfected cells will be puromycin resistant, however, only if growth of Runx1-1 (or Luc) KD cells is permissible GFP+ cells are observed.

2.5.12 Western Blot

Methods were described previously (Osorio *et al.*, 2008), but here we added phosphatase inhibitor cocktail (Thermo Fisher). K14CreERT2; $Runx1^{fl/fl}$ cells were treated either with ethanol (WT) or 4-OHT for 4 or 5 days. OSM was added when required on the last day. The following antibodies were used in TBS-T buffer with 5% BSA: rabbit: [pStat3-T703 (1:2000, Cell Signaling #9131), pStat3-S727 (1:1000, Cell Signaling #9134), Stat3 (1:1000, Cell Signaling #9132), Runx1-31 (1:3000, SDIX generated No31)]; mouse: [actin-C4 (1:8000, Millipore #mab1501)].

2.5.13 QRT-PCR

RNA from iKO cells treated with ethanol or 4-OHT (Sigma) for 4 days was extracted using Trizol and the RNeasy Mini kit (Qiagen) and converted to cDNA with the iScript kit (Biorad). The mouse Jak/Stat RT2 Profiler PCR array (PAMM-039) was purchased from Qiagen, used as instructed, and the data was analyzed using proprietary bioinformatics tools provided by Qiagen. Details for Runx1 QRT-PCR were described previously (Hoi *et al.*, 2010).

2.5.14 ChIP

~0.5 - 1 x 10^7 wild-type ($Runx1^{fl/fl}$) and inducible Runx1 knockout (iKO) (K14-CreERT2 x $Runx1^{fl/fl}$) keratinocytes were treated with 4-OHT for 4 days prior to fixation. All steps were carried out according to (Ortt and Sinha, 2010) with minor variations. Cells were fixed for 3 minutes. Sonic Dismembrator model 100 (Fisher Scientific) was used to sonicate the chromatin and protein A agarose beads (Millipore) for immunoprecipitation. SOCS3 and SOCS4 genomic regions were pulled down with 5μg Runx1 antibody (Abcam ab23980). Control ChIP reactions were setup with 4μg rabbit-IgG (Imgenex 20304) and 3μg H3 (abcam ab1791). ChIP DNA was purified with the PCR purification kit (Qiagen). ChIP-QPCR signal of Runx1 and IgG were interpolated from serial dilution of input DNA standards. Next, fold enrichment of Runx1 over IgG signal was compared between different promoter regions in WT and iKO samples.

Primers: SOCS3 [1 (F:5'- ccaagcagcccagccagagt-3', R:5'- tggcttcaaagggcttgcaggg-3'); 2 [(F:5'- tgggtcggggaacagacagca -3', R:5'- caacccagatggctcaccgca -3')]; SOCS4 [1 (F:5'- caggctgtgaacaccagtgtcagg -3', R:5'- cagcgagatgaggtagtcggcct -3'); 2(F:5'- actctgggctgaccgtggagc -3', R:5'- cgctcgctcaagcgcggaa -3'); 3(F:5'- gagctcctcgggaccgggtt -3', R:5'- ttgagggcacgaagggcgtg -3'); 4(F:5'- tgtgatccctgtgagaggtagcc -3', R:5'- gcgagatgaggtagtcggcctgt -3')]. Gapdh(F:5'- accgaagaacaacgaggaga -3', R:5'- gagaacaggggaaatggaga -3'')].

2.5.15 Microscopy and Image processing

Images were acquired using the IP-Lab software (MVI) on a light fluorescence microscope (Nikon) equipped with a CCD 12-bit digital camera (Retiga EXi, QImaging) and motorized z-stage. Single images were color combined and were enhanced for brightness, contrast and levels and assembled in montages using Adobe Photoshop and Illustrator.

2.5.16 Statistics

Data was processed and analyzed in Excel or R. The survival package in R was used for Kaplan-Meier analysis.

REFERENCES

Abel EL, Angel JM, Kiguchi K & DiGiovanni J (2009) Multi-stage chemical carcinogenesis in mouse skin: fundamentals and applications. *Nature protocols* **4:** 1350-1362

Alam M & Ratner D (2001) Cutaneous Squamous Cell Carcinoma. *New England Journal of Medicine* **344:** 975-983

Andersson A, Ritz C, Lindgren D, Edén P, Lassen C, Heldrup J, Olofsson T, Råde J, Fontes M, Porwit-Macdonald A, Behrendtz M, Höglund M, Johansson B & Fioretos T (2007) Microarray-based classification of a consecutive series of 121 childhood acute leukemias: prediction of leukemic and genetic subtype as well as of minimal residual disease status. *Leukemia: official journal of the Leukemia Society of America, Leukemia Research Fund, U.K* **21:** 1198-1203

Ang Y-S, Tsai S-Y, Lee D-F, Monk J, Su J, Ratnakumar K, Ding J, Ge Y, Darr H, Chang B, Wang J, Rendl M, Bernstein E, Schaniel C & Lemischka IR (2011) Wdr5 mediates self-renewal and reprogramming via the embryonic stem cell core transcriptional network. *Cell* **145:** 183-197

Appleford PJ & Woollard A (2009) RUNX genes find a niche in stem cell biology. *Journal of cellular biochemistry* **108:** 14-21

Aziz MH, Manoharan HT, Sand JM & Verma AK (2007) Protein Kinase Ce Interacts With Stat3 and Regulates Its Activation That Is Essential for the Development of Skin Cancer. *Molecular Carcinogenesis* **653:** 646-653

Badea L, Herlea V, Dima SO, Dumitrascu T & Popescu I Combined gene expression analysis of whole-tissue and microdissected pancreatic ductal adenocarcinoma identifies genes specifically overexpressed in tumor epithelia. *Hepato-gastroenterology* **55:** 2016-2027

Barker N, van de Wetering M & Clevers H (2008) The intestinal stem cell. *Genes & development* **22:** 1856-1864

Beroukhim R, Brunet J-P, Di Napoli A, Mertz KD, Seeley A, Pires MM, Linhart D, Worrell RA, Moch H, Rubin MA, Sellers WR, Meyerson M, Linehan WM, Kaelin WG & Signoretti S (2009) Patterns of gene expression and copy-number alterations in von-hippel lindau disease-associated and sporadic clear cell carcinoma of the kidney. *Cancer research* **69:** 4674-4681

Bickenbach JR & Mackenzie IC (1984) Identification and localization of label-retaining cells in hamster epithelia. *The Journal of investigative dermatology* **82:** 618-622

Bito T, Sumita N, Ashida M, Budiyanto A, Ueda M, Ichihashi M, Tokura Y & Nishigori C (2011) Inhibition of Epidermal Growth Factor Receptor and PI3K/Akt Signaling Suppresses Cell Proliferation and Survival through Regulation of Stat3 Activation in Human Cutaneous Squamous Cell Carcinoma. *Journal of skin cancer* **2011:** 874571

Blyth K, Cameron ER & Neil JC (2005) The RUNX genes: gain or loss of function in cancer. *Nature reviews. Cancer* **5:** 376-3787

Bredel M, Bredel C, Juric D, Harsh GR, Vogel H, Recht LD & Sikic BI (2005) Functional network analysis reveals extended gliomagenesis pathway maps and three novel MYC-interacting genes in human gliomas. *Cancer research* **65:** 8679-8689

Bromberg JF, Wrzeszczynska MH, Devgan G, Zhao Y, Pestell RG, Albanese C & Darnell JE (1999) Stat3 as an oncogene. *Cell* **98:** 295-303

Casanova ML, Bravo A, Martínez-Palacio J, Fernández-Aceñero MJ, Villanueva C, Larcher F, Conti CJ & Jorcano JL (2004) Epidermal abnormalities and increased malignancy of skin tumors in human epidermal keratin 8-expressing transgenic mice. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* **18:** 1556-1558

Chan KS, Carbajal S, Kiguchi K, Clifford J, Sano S & DiGiovanni J (2004a) Epidermal Growth Factor Receptor-Mediated Activation of Stat3 during Multistage Skin Carcinogenesis. *Cancer Research* **64:** 2382-2389

Chan KS, Sano S, Kataoka K, Abel E, Carbajal S, Beltran L, Clifford J, Peavey M, Shen J & Digiovanni J (2008) Forced expression of a constitutively active form of Stat3 in mouse epidermis enhances malignant progression of skin tumors induced by two-stage carcinogenesis. *Oncogene* **27:** 1087-1094

Chan KS, Sano S, Kiguchi K, Anders J, Komazawa N, Takeda J & Digiovanni J (2004b) Disruption of Stat3 reveals a critical role in both the initiation and the promotion stages of epithelial carcinogenesis. *Journal of Clinical Investigation* **114:** 720-728

Chuang C-H, Wallace MD, Abratte C, Southard T & Schimenti JC (2010) Incremental genetic perturbations to MCM2-7 expression and subcellular

distribution reveal exquisite sensitivity of mice to DNA replication stress. *PLoS genetics* **6:**

Crabtree JS, Jelinsky SA, Harris HA, Choe SE, Cotreau MM, Kimberland ML, Wilson E, Saraf KA, Liu W, McCampbell AS, Dave B, Broaddus RR, Brown EL, Kao W, Skotnicki JS, Abou-Gharbia M, Winneker RC & Walker CL (2009) Comparison of human and rat uterine leiomyomata: identification of a dysregulated mammalian target of rapamycin pathway. *Cancer research* **69:** 6171-6178

Detwiller KY, Fernando NT, Segal NH, Ryeom SW, D'Amore PA & Yoon SS (2005) Analysis of hypoxia-related gene expression in sarcomas and effect of hypoxia on RNA interference of vascular endothelial cell growth factor A. *Cancer research* **65:** 5881-5889

Douglas WG, Tracy E, Tan D, Yu J, Jr WLH, Rigual NR, Loree TR, Wang Y & Baumann H (2004) Development of Head and Neck Squamous Cell Carcinoma Is Associated With Altered Cytokine Responsiveness. *Cancer Res* **8:** 585-593

Friedman AD (2009) Cell cycle and developmental control of hematopoiesis by Runx1. *Journal of cellular physiology* **219:** 520-524

Frierson HF, El-Naggar AK, Welsh JB, Sapinoso LM, Su AI, Cheng J, Saku T, Moskaluk CA & Hampton GM (2002) Large scale molecular analysis identifies genes with altered expression in salivary adenoid cystic carcinoma. *The American journal of pathology* **161:** 1315-1323

Gartsbein M, Alt A, Hashimoto K, Nakajima K, Kuroki T & Tennenbaum T (2006) The role of protein kinase C delta activation and STAT3 Ser727 phosphorylation in insulin-induced keratinocyte proliferation. *Journal of cell science* **119:** 470-481

Gilad O, Nabet BY, Ragland RL, Schoppy DW, Smith KD, Durham AC & Brown EJ (2010) Combining ATR suppression with oncogenic Ras synergistically increases genomic instability, causing synthetic lethality or tumorigenesis in a dosage-dependent manner. *Cancer research* **70:** 9693-9702

Ginos MA, Page GP, Michalowicz BS, Patel KJ, Volker SE, Pambuccian SE, Ondrey FG, Adams GL & Gaffney PM (2004) Identification of a gene expression signature associated with recurrent disease in squamous cell carcinoma of the head and neck. *Cancer research* **64:** 55-63

Giordano TJ, Au AYM, Kuick R, Thomas DG, Rhodes DR, Wilhelm KG, Vinco M, Misek DE, Sanders D, Zhu Z, Ciampi R, Hanash S, Chinnaiyan A, Clifton-Bligh RJ, Robinson BG, Nikiforov YE & Koenig RJ (2006) Delineation, functional validation, and bioinformatic evaluation of gene expression in thyroid follicular carcinomas with the PAX8-PPARG translocation. *Clinical cancer research* **12:** 1983-1993

Growney JD, Shigematsu H, Li Z, Lee BH, Adelsperger J, Rowan R, Curley DP, Kutok JL, Akashi K, Williams IR, Speck N & Gilliland DG (2005) Loss of Runx1 perturbs adult hematopoiesis and is associated with a myeloproliferative phenotype. *Blood* **106:** 494-504

Gumz ML, Zou H, Kreinest PA, Childs AC, Belmonte LS, LeGrand SN, Wu KJ, Luxon BA, Sinha M, Parker AS, Sun L-Z, Ahlquist DA, Wood CG & Copland JA (2007) Secreted frizzled-related protein 1 loss contributes to tumor phenotype of clear cell renal cell carcinoma. *Clinical cancer research* **13:** 4740-4749

Haferlach T, Kohlmann A, Wieczorek L, Basso G, Kronnie GT, Béné M-C, De Vos J, Hernández JM, Hofmann W-K, Mills KI, Gilkes A, Chiaretti S, Shurtleff SA, Kipps TJ, Rassenti LZ, Yeoh AE, Papenhausen PR, Liu W-M, Williams PM & Foà R (2010) Clinical utility of microarray-based gene expression profiling in the diagnosis and subclassification of leukemia: report from the International Microarray Innovations in Leukemia Study Group. *Journal of clinical oncology* **28:** 2529-2537

Hao Y, Triadafilopoulos G, Sahbaie P, Young HS, Omary MB & Lowe AW (2006) Gene expression profiling reveals stromal genes expressed in common between Barrett's esophagus and adenocarcinoma. *Gastroenterology* **131:** 925-933

Hoi C, Lee SE, Lu S-Y, McDermitt DJ, Osorio KM, Piskun C, Peters R, Paus R & Tumbar T (2010) Runx1 directly promotes proliferation of hair follicle stem cells and epithelial tumor formation in mouse skin. *Molecular and cellular biology* **30:** 2518-2536

Hong Y, Downey T, Eu KW, Koh PK & Cheah PY (2010) A "metastasis-prone" signature for early-stage mismatch-repair proficient sporadic colorectal cancer patients and its implications for possible therapeutics. *Clinical & experimental metastasis* **27:** 83-90

Huang S-P, Lan Y-H, Lu T-L, Pao J-B, Chang T-Y, Lee H-Z, Yang W-H, Hsieh C-J, Chen L-M, Huang L-C, Ting W-C & Bao B-Y (2011) Clinical significance of

runt-related transcription factor 1 polymorphism in prostate cancer. *BJU international* **107:** 486-492

Ihle JN (1996) STATs: Signal Transducers and Activators of Transcription. *Cell* **84:** 331-334

Ishihara K & Hirano T (2002) Molecular basis of the cell specificity of cytokine action. *Biochimica et biophysica acta* **1592:** 281-296

Jackson EL, Willis N, Mercer K, Bronson RT, Crowley D, Montoya R, Jacks T & Tuveson DA (2001) Analysis of lung tumor initiation and progression using conditional expression of oncogenic K-ras. *Genes & development* **15:** 3243-3248

Kim DJ, Kataoka K, Rao D, Kiguchi K, Cotsarelis G & Digiovanni J (2009) Targeted disruption of stat3 reveals a major role for follicular stem cells in skin tumor initiation. *Cancer research* **69:** 7587-7594

Kopan R, Lee J, Lin M-H, Syder AJ, Kesterson J, Crutchfield N, Li CR, Wu W, Books J & Gordon JI (2002) Genetic mosaic analysis indicates that the bulb region of coat hair follicles contains a resident population of several active multipotent epithelial lineage progenitors. *Developmental biology* **242:** 44-57

Korkola JE, Houldsworth J, Chadalavada RSV, Olshen AB, Dobrzynski D, Reuter VE, Bosl GJ & Chaganti RSK (2006) Down-regulation of stem cell genes, including those in a 200-kb gene cluster at 12p13.31, is associated with in vivo differentiation of human male germ cell tumors. *Cancer research* **66:** 820-827

Lapouge G, Youssef KK, Vokaer B, Achouri Y, Michaux C, Sotiropoulou P a & Blanpain C (2011) Identifying the cellular origin of squamous skin tumors. *Proceedings of the National Academy of Sciences of the United States of America* **108:** 7431-7436

Larsen L & Ropke C (2002) Suppressors of cytokine signalling: SOCS. Review article. *Apmis* **110:** 833-844

Lee J, Kotliarova S, Kotliarov Y, Li A, Su Q, Donin NM, Pastorino S, Purow BW, Christopher N, Zhang W, Park JK & Fine HA (2006) Tumor stem cells derived from glioblastomas cultured in bFGF and EGF more closely mirror the phenotype and genotype of primary tumors than do serum-cultured cell lines. *Cancer cell* **9:** 391-403

Li N, Grivennikov SI & Karin M (2011) The unholy trinity: inflammation, cytokines, and STAT3 shape the cancer microenvironment. *Cancer cell* **19:** 429-431

Logsdon CD, Simeone DM, Binkley C, Arumugam T, Greenson JK, Giordano TJ, Misek DE, Kuick R & Hanash S (2003) Molecular profiling of pancreatic adenocarcinoma and chronic pancreatitis identifies multiple genes differentially regulated in pancreatic cancer. *Cancer research* **63:** 2649-2657

Madisen L, Zwingman TA, Sunkin SM, Oh SW, Zariwala HA, Gu H, Ng LL, Palmiter RD, Hawrylycz MJ, Jones AR, Lein ES & Zeng H (2010) A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nature neuroscience* **13:** 133-140

Magee JA, Piskounova E & Morrison SJ (2012) Cancer Stem Cells: Impact, Heterogeneity, and Uncertainty. *Cancer Cell* **21**: 283–296

Malanchi I, Peinado H, Kassen D, Hussenet T, Metzger D, Chambon P, Huber M, Hohl D, Cano A, Birchmeier W & Huelsken J (2008) Cutaneous cancer stem cell maintenance is dependent on beta-catenin signalling. *Nature* **452:** 650-653

Mangan JK & Speck NA (2011) RUNX1 mutations in clonal myeloid disorders: from conventional cytogenetics to next generation sequencing, a story 40 years in the making. *Critical reviews in oncogenesis* **16:** 77-91

McLean IW & Nakane PK (1974) Periodate-lysine-paraformaldehyde fixative. A new fixation for immunoelectron microscopy. *The journal of histochemistry and cytochemistry : official journal of the Histochemistry Society* **22:** 1077-1083

Mikhail FM, Sinha KK, Saunthararajah Y & Nucifora G (2006) Normal and transforming functions of RUNX1: a perspective. *Journal of cellular physiology* **207:** 582-593

Molinolo AA, Amornphimoltham P, Squarize CH, Castilho RM, Patel V & Gutkind JS (2009) Dysregulated molecular networks in head and neck carcinogenesis. *Oral oncology* **45:** 324-334

North T, Gu T, Stacy T, Wang Q, Howard L, Binder M, Marin-Padilla M & Speck N (1999) Cbfa2 is required for the formation of intra-aortic hematopoietic clusters. *Development* **126:** 2563-2575

Ortt K & Sinha S (2010) Chromatin immunoprecipitation for identifying transcription factor targets in keratinocytes. *Methods in molecular biology* **585:** 159-170

Osorio KM, Lee SE, McDermitt DJ, Waghmare SK, Zhang YV, Woo HN & Tumbar T (2008) Runx1 modulates developmental, but not injury-driven, hair follicle stem cell activation. *Development* **135:** 1059-68

Osorio KM, Lilja KC & Tumbar T (2011) Runx1 modulates adult hair follicle stem cell emergence and maintenance from distinct embryonic skin compartments. *The Journal of cell biology* **193:** 235-250

Pierceall WE, Goldberg LH, Tainsky MA, Mukhopadhyay T & Ananthaswamy HN (1991) Ras gene mutation and amplification in human nonmelanoma skin cancers. *Molecular carcinogenesis* **4:** 196-202

Planagumà J, Gonzalez M, Doll A, Monge M, Gil-Moreno A, Baró T, García A, Xercavins J, Alameda F, Abal M & Reventós J (2006) The up-regulation profiles of p21WAF1/CIP1 and RUNX1/AML1 correlate with myometrial infiltration in endometrioid endometrial carcinoma. *Human pathology* **37:** 1050-1057

Planagumà J, Liljeström M, Alameda F, Bützow R, Virtanen I, Reventós J, Hukkanen M & Bützow R (2011) Matrix metalloproteinase-2 and matrix metalloproteinase-9 codistribute with transcription factors RUNX1/AML1 and ETV5/ERM at the invasive front of endometrial and ovarian carcinoma. *Human pathology* **42:** 57-67

Pyeon D, Newton MA, Lambert PF, den Boon JA, Sengupta S, Marsit CJ, Woodworth CD, Connor JP, Haugen TH, Smith EM, Kelsey KT, Turek LP & Ahlquist P (2007) Fundamental differences in cell cycle deregulation in human papillomavirus-positive and human papillomavirus-negative head/neck and cervical cancers. *Cancer research* **67:** 4605-4619

Quade BJ, Wang T-Y, Sornberger K, Dal Cin P, Mutter GL & Morton CC (2004) Molecular pathogenesis of uterine smooth muscle tumors from transcriptional profiling. *Genes, chromosomes & cancer* **40:** 97-108

Rheinwald JG & Beckett MA (1980) Defective terminal differentiation in culture as a consistent and selectable character of malignant human keratinocytes. *Cell* **22:** 629–632

Rheinwald JG & Beckett MA (1981) Tumorigenic keratinocyte lines requiring anchorage and fibroblast support cultured from human squamous cell carcinomas. *Cancer research* **41:** 1657

Rho O, Kim DJ, Kiguchi K & Digiovanni J (2011) Growth factor signaling pathways as targets for prevention of epithelial carcinogenesis. *Molecular carcinogenesis* **50:** 264-279

Riker AI, Enkemann SA, Fodstad O, Liu S, Ren S, Morris C, Xi Y, Howell P, Metge B, Samant RS, Shevde LA, Li W, Eschrich S, Daud A, Ju J & Matta J (2008) The gene expression profiles of primary and metastatic melanoma yields a transition point of tumor progression and metastasis. *BMC medical genomics* **1:**

Sanchez-Carbayo M, Socci ND, Lozano J, Saint F & Cordon-Cardo C (2006) Defining molecular profiles of poor outcome in patients with invasive bladder cancer using oligonucleotide microarrays. *Journal of clinical oncology* **24:** 778-789

Samokhvalov IM, Samokhvalova NI & Nishikawa S-ichi (2007) Cell tracing shows the contribution of the yolk sac to adult haematopoiesis. *Nature* **446:** 1056-1061

Sano S, Chan KS, Carbajal S, Clifford J, Peavey M, Kiguchi K, Itami S, Nickoloff BJ & DiGiovanni J (2005) Stat3 links activated keratinocytes and immunocytes required for development of psoriasis in a novel transgenic mouse model. *Nature medicine* **11:** 43-49

Sano S, Itami S, Takeda K, Tarutani M, Yamaguchi Y, Miura H, Yoshikawa K, Akira S & Takeda J (1999) Keratinocyte-specific ablation of Stat3 exhibits impaired skin remodeling, but does not affect skin morphogenesis. *The EMBO journal* **18:** 4657-4668

Sano S, Kira M, Takagi S, Yoshikawa K, Takeda J & Itami S (2000) Two distinct signaling pathways in hair cycle induction: Stat3-dependent and -independent pathways. *Proceedings of the National Academy of Sciences of the United States of America* **97:** 13824-1389

Santos M, Ballestín C, Garcia-Martín R & Jorcano JL (1997) Delays in malignant tumor development in transgenic mice by forced epidermal keratin 10 expression in mouse skin carcinomas. *Molecular carcinogenesis* **20:** 3-9

Segara D, Biankin AV, Kench JG, Langusch CC, Dawson AC, Skalicky DA, Gotley DC, Coleman MJ, Sutherland RL & Henshall SM (2005) Expression of HOXB2, a retinoic acid signaling target in pancreatic cancer and pancreatic intraepithelial neoplasia. *Clinical cancer research* **11:** 3587-3596

Segatto O, Anastasi S & Alemà S (2011) Regulation of epidermal growth factor receptor signalling by inducible feedback inhibitors. *Journal of cell science* **124:** 1785-1793

Sikandar SS, Pate KT, Anderson S, Dizon D, Edwards RA, Waterman ML & Lipkin SM (2010) NOTCH signaling is required for formation and self-renewal of tumor-initiating cells and for repression of secretory cell differentiation in colon cancer. *Cancer research* **70**: 1469–1478

Slattery ML, Lundgreen A, Herrick JS, Caan BJ, Potter JD & Wolff RK (2011) Associations between genetic variation in RUNX1, RUNX2, RUNX3, MAPK1 and eIF4E and riskof colon and rectal cancer: additional support for a TGF-β-signaling pathway. *Carcinogenesis* **32:** 318-326

Soriano P (1999) Generalized lacZ expression with the ROSA26 Cre reporter strain. *Nature genetics* **21:** 70-71

Spencer JM, Kahn SM, Jiang W, DeLeo VA & Weinstein IB (1995) Activated ras genes occur in human actinic keratoses, premalignant precursors to squamous cell carcinomas. *Archives of dermatology* **131:** 796-800

Stegmaier K, Ross KN, Colavito SA, O'Malley S, Stockwell BR & Golub TR (2004) Gene expression-based high-throughput screening(GE-HTS) and application to leukemia differentiation. Nature genetics 36: 257-263

Su L-J, Chang C-W, Wu Y-C, Chen K-C, Lin C-J, Liang S-C, Lin C-H, Whang-Peng J, Hsu S-L, Chen C-H & Huang C-YF (2007) Selection of DDX5 as a novel internal control for Q-RT-PCR from microarray data using a block bootstrap re-sampling scheme. BMC genomics 8: 140

Sun L, Hui A-M, Su Q, Vortmeyer A, Kotliarov Y, Pastorino S, Passaniti A, Menon J, Walling J, Bailey R, Rosenblum M, Mikkelsen T & Fine HA (2006) Neuronal and glioma-derived stem cell factor induces angiogenesis within the brain. Cancer cell 9: 287-300

Tanaka S (2009) [Biological roles of signaling adaptor protein CRK]. *Seikagaku. The Journal of Japanese Biochemical Society* **81:** 361-376

TCGA (2011) No Associated Paper.

Tumbar T (2006) Epithelial skin stem cells. *Methods in enzymology* **419:** 73-99

Tumbar T (2012) Ontogeny and Homeostasis of Adult Epithelial Skin Stem Cells. *Stem cell rev* **8**: 561-576

Visvader JE (2011) Cells of origin in cancer. *Nature* **469:** 314-322

Vitale-Cross L, Amornphimoltham P, Fisher G, Molinolo AA & Gutkind JS (2004) Conditional expression of K-ras in an epithelial compartment that includes the stem cells is sufficient to promote squamous cell carcinogenesis. *Cancer research* **64:** 8804-8807

Wang GY, Wang J, Mancianti M-L & Epstein Jr. EH (2011) Basal Cell Carcinomas Arise from Hair Follicle Stem Cells in Ptch1+/− Mice. *Cancer Cell* **19:** 114-124

White AC, Tran K, Khuu J, Dang C, Cui Y, Binder SW & Lowry WE (2011) Defining the origins of Ras/p53-mediated squamous cell carcinoma. *Proceedings of the National Academy of Sciences of the United States of America* **108:** 7425-7430

White JS, Weissfeld JL, Ragin CCR, Rossie KM, Martin CL, Shuster M, Ishwad CS, Law JC, Myers EN, Johnson JT & Gollin SM (2007) The influence of clinical and demographic risk factors on the establishment of head and neck squamous cell carcinoma cell lines. *Oral oncology* **43:** 701-612

Wilson NK, Foster SD, Wang X, Knezevic K, Schütte J, Kaimakis P, Chilarska PM, Kinston S, Ouwehand WH, Dzierzak E, Pimanda JE, de Bruijn MFTR & Göttgens B (2010) Combinatorial transcriptional control in blood stem/progenitor cells: genome-wide analysis of ten major transcriptional regulators. *Cell stem cell* **7:** 532-544

Yeh H-Y, Cheng S-W, Lin Y-C, Yeh C-Y, Lin S-F & Soo V-W (2009) Identifying significant genetic regulatory networks in the prostate cancer from microarray data based on transcription factor analysis and conditional independency. *BMC medical genomics* **2:** 70

Yokogami K, Wakisaka S, Avruch J & Reeves S a (2000) Serine phosphorylation and maximal activation of STAT3 during CNTF signaling is mediated by the rapamycin target mTOR. *Current biology : CB* **10:** 47-50

Youssef KK, Van Keymeulen A, Lapouge G, Beck B, Michaux C, Achouri Y, Sotiropoulou P a & Blanpain C (2010) Identification of the cell lineage at the origin of basal cell carcinoma. *Nature cell biology* **12:** 299-305

Yu H & Jove R (2004) The STATs of cancer--new molecular targets come of age. *Nature reviews. Cancer* **4:** 97-105

Yusenko MV, Kuiper RP, Boethe T, Ljungberg B, van Kessel AG & Kovacs G (2009) High-resolution DNA copy number and gene expression analyses

distinguish chromophobe renal cell carcinomas and renal oncocytomas. *BMC cancer* **9:** 152

Yuspa SH, Hawley-Nelson P, Koehler B & Stanley JR (1980) A survey of transformation markers in differentiating epidermal cell lines in culture. *Cancer research* **40:** 4694-4703

Zhan F, Barlogie B, Arzoumanian V, Huang Y, Williams DR, Hollmig K, Pineda-Roman M, Tricot G, van Rhee F, Zangari M, Dhodapkar M & Shaughnessy JD (2007) Gene-expression signature of benign monoclonal gammopathy evident in multiple myeloma is linked to good prognosis. *Blood* **109:** 1692-1700

Zhang YV, Cheong J, Ciapurin N, McDermitt DJ & Tumbar T (2009) Distinct Self-Renewal and Differentiation Phases in the Niche of Infrequently Dividing Hair Follicle Stem Cells. *Cell stem cell* **5:** 267-278

CHAPTER 3[3]

HERITABILITY AND INTER-POPULATION DIFFERENCES IN LIPID PROFILES OF DROSOPHILA MELANOGASTER

3.1 Overview

Characterizing and understanding the complex spectrum of lipids in higher organisms lags far behind our analysis of genome and transcriptome sequences. Here we generate and evaluate comprehensive lipid profiles (>200 lipids) of 92 inbred lines from 5 different *Drosophila melanogaster* populations. We find that the majority of lipid species are highly heritable, and lipids with odd-chain fatty acids, which cannot be generated by the fly itself, have at least equal heritabilities. Abundance of the endosymbiont *Wolbachia*, a potential provider of odd-chained lipids, was positively correlated with this group of lipids. Additionally, we show that despite years of laboratory rearing on the same medium, the lipid profiles of the five geographic populations are sufficiently distinct for population discrimination. Our data predict a strikingly different membrane fluidity for Netherlands flies, which is supported by their increased ethanol tolerance. We find that 18.5% of lipids species showing strong high/low concentration differences between males and females. Through an analysis of the

[3] This work is in submission to G3 "Scheitz CJS , Guo Y, Harshman LG, Clark AG(2012) Heritability and inter-population differences in lipid profiles of *Drosophila melanogaster*. 3G". The author contributions are as follows: Cornelia Scheitz designed the project, performed the lipid extractions and analyzed and interpreted all experiments, created all figures and wrote the manuscript. Yu Guo assisted in DAPC analysis and network computation. Lawrence Harshman and Andrew Clark assisted in project design, data interpretation and manuscript preparation.

correlation structure of the lipid classes, we find modules of co-regulated lipids, and begin to associate these with metabolic constraints. Our data provide a foundation for developing associations between variation in lipid composition with variation in the genome as well as with metrics of health and overall reproductive fitness.

3.2 Introduction

Drosophila melanogaster is such an established and powerful model organisms for a range of biological investigation, from fundamental cell biology to population genetics. Thus analysis of a richly complex set of biochemical traits is likely to progress rapidly (Consortium, 2007; Sowell et al., 2007; Graveley et al., 2011; Kharchenko et al., 2011). The study of lipid biology of *Drosophila* also has a long history (Taraschi and Rubin, 1985), but we have yet to determine a more global picture of variation in lipid composition, the genetic means by which this variation is regulated, nor its influence on cellular functions.

One area of lipid biology that has been well characterized in ectotherms, including *D. melanogaster,* is homeoviscous adaptation (Sinensky, 1974; McElhaney and Souza, 1976; Cossins and Prosser, 1978; Hazel and Williams, 1990; Hazel, 1995). This refers to the global adaptation of membrane lipid content after temperature changes by altering the ratio of phosphatidylcholines over phosphatidylethanolamines (PC/PE) and the amount of FA unsaturation. Both are crucial for acclimation by controlling membrane fluidity (Hazel, 1995; Tiku et al., 1996; Cossins et al., 2002; Overgaard et al., 2008; Cooper et al., 2011;

Koštál et al., 2011). There is a baseline ability to acclimate to changes in temperature (Ohtsu et al., 1998; Overgaard et al., 2008), and there is also evidence that selection acts on this ability, particularly with respect to cold-acclimation (Cooper et al., 2011). While the environmental contributions to acclimation and membrane composition have been studied, the heritability of these lipid components has not been analyzed to date. Additionally, there is a lack of knowledge about the biological consequences of lipid variation. For instance, phosphatidylserines (PS) are important for the recognition and clearance of apoptotic cells (Vance, 2003), but there is no report of the role of fatty acid (FA) chain length or desaturation on the efficacy of PS on apoptosis. The same is true for the response to phosphatidylinositols (PI), which induce the PIP3-kinase cascade playing vital roles in cell growth and survival (Cully et al., 2006).

Here we present the most comprehensive analysis of lipids across populations of *D. melanogaster* to date and for the first time analyze the heritability of each component across 5 distinct populations. The 233 lipids analyzed span all major lipid classes with FAs of various length (odd and even carbon chain lengths) and include different degrees of FA saturation. In this dataset, we detect heritable differences in membrane fluidity and, related to it, ethanol tolerance across populations. We also correlate the lipid profiles with the abundance of the intra-cellular microbe *Wolbachia*. This analysis points toward a trade-off between gaining odd-chained lipids and being at risk for apoptosis. Additionally, we investigate the structure underlying the lipid network in males and females by

analyzing the modular pattern of clustering of correlations and using discriminant analysis of principal components (DAPC). We find striking differences that point toward the existence of an additional layer of lipid network interactions in females.

3.3 Results

We isolated total lipid from whole adult flies of 92 inbred lines belonging to 5 populations. For each line we extracted lipids from three replicates of groups of 10 adults of each sex (separately) using the Bligh and Dryer method. Lipid species were identified using a mass spectrometry method described below. From the initial 342 lipids 233 were maintained for further analysis while the remaining lipids did not pass quality control (see Materials and Methods). Our dataset includes PEs (33 species), PCs (34 species), PSs (12 species), PIs (18 species), phosphatidylglycerols (PGs – 22 species), phosphatidic acids (PAs – 2 species), lysophillic PCs (9 species) and PEs (8 species) as well as triacylglycerols (TAGs – 61 species) and diacylglycerols (DAGs – 33 species) (see Appendix 3). The excluded lipids belonged to all classes (Figure 3-1A). Odd and even carbon-chain fatty acids (FAs) were excluded at roughly equal frequencies (Figure 3-1B). We find that >50% of lipids with FAs of length $C_{20\text{-}22}$ doe not pass quality control, but this percentage is much lower in all other bins (Figure 3-1C). A recent report of lipid profiles in *D. melanogaster* by Hammad et al. (2011) showed absence of FAs greater than $C_{18.}$ However, while C_{16} and C_{18} are also in our dataset the most

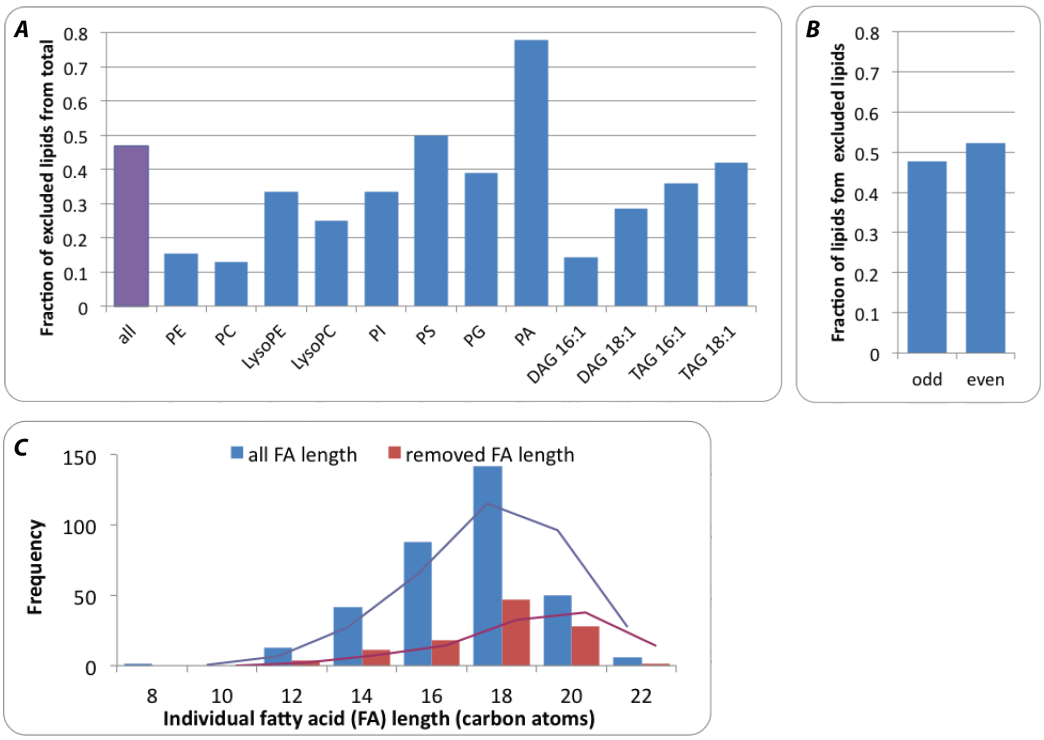


Figure 3-1: Quality control of lipid profiles

A: Fraction of lipid species that were excluded in quality control because they showed either no concentration across all lines or were not normally distributed. X-axis indicates the respective total that each bar was calculated in relation to. B: Same as A but grouped by FA length with the categories odd and even. C: Frequency spectrum of estimated individual FA length. Lines are moving averages to help visualize the trans in the spectra.

common FAs, we do detect lipid species containing C_{20} and C_{22} FAs (Figure 3-1C) similar to what is detected in other studies (Koštál et al., 2011; Parisi et al., 2011).

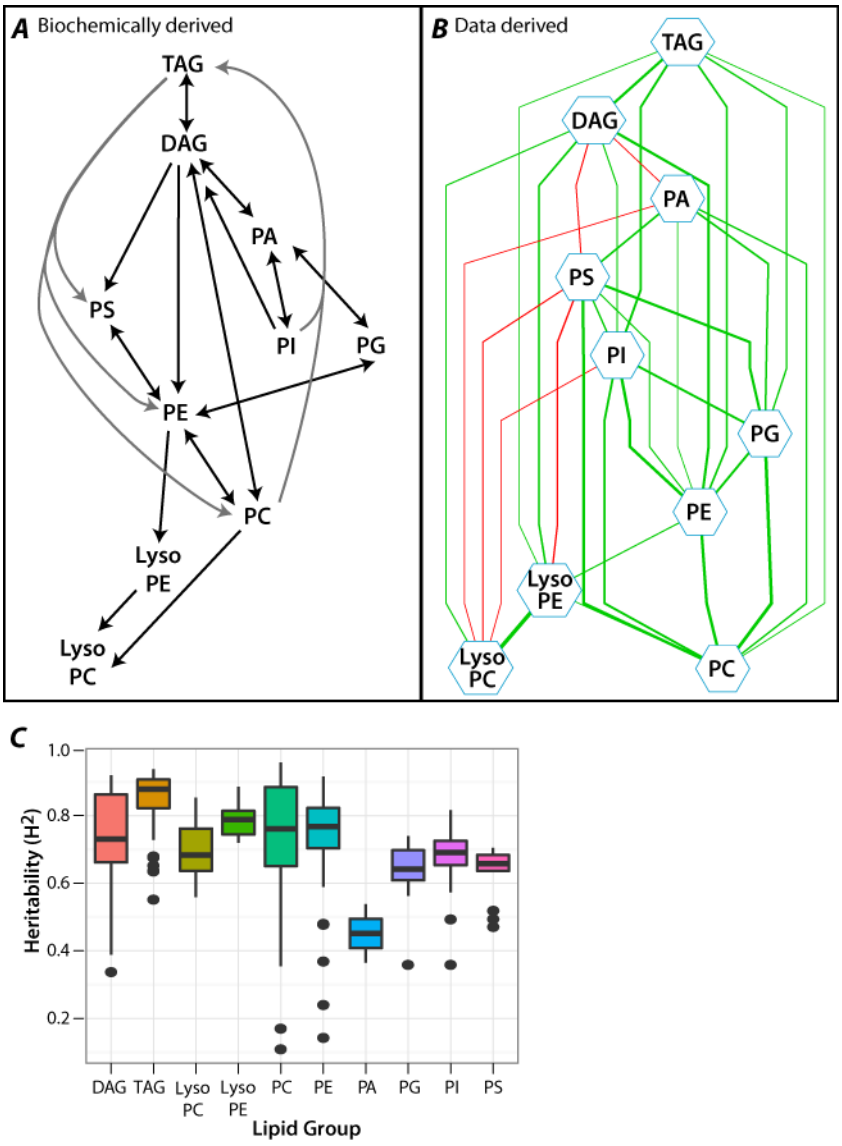
3.3.1 Lipids fall within the established top-level network and are highly heritable

First we wanted to confirm the quality of our dataset by comparing a data-derived top level network to known biochemical pathways (Coleman, 2004 and Figure Figure 3-2A). After removing population and sex effects with linear mixed models (REML) we calculated pair-wise Pearson Correlation Coefficients (PCC) between all lipid classes and constructed a network view of the relationships among these lipid classes using Cytoscape (Cline et al., 2007 and Figure 3-2B). All top-level correlations greater than |0.1| automatically arranged by their cluster distance are displayed and this network shows broad consistency with the published literature, however, our data suggest additional novel links that remain to be validated biochemically.

Because our measurements were done in a collection of highly inbred lines, we were able to estimate the heritability of the lipids in this network. We calculated the broad-sense heritability (H^2) for each lipid (plotted by lipid class Figure 3-2C) by considering variance in line means over the total phenotypic variance (using linear models estimated by REML). While there is a high level of heritability across lipid classes, and, with the exception of PAs, the different classes display a considerable range of genetic contribution (Figure 3-2C). This is the first evidence that lipid composition in *D. melanogaster* is to a large extent regulated

Figure 3-2: Lipid data fall into known top-level network and has high heritability

A: Top-level network derived from the literature (Coleman, 2004). B: Top-level lipid network constructed with Cytoscape. All pairs of lipid classes with PCC > 0.1 or PCC < -0.1 are connected in the network. A positive correlation between two lipid classes is represented by a green edge, while a negative correlation is represented by a red edge. The thickness of the edges is proportional to the magnitude of PCC between the lipid classes. C: Broad sense heritability (H2) for each measured lipid species grouped by lipid class.


A Biochemically derived
TAG
DAG
PA
PS
PI
PG
PE
PC
Lyso PE
Lyso PC
B Data derived
TAG
DAG
PA
PS
PI
PG
PE
Lyso PE
Lyso PC
PC
C
Heritability (H2)
1.0
0.8
0.6
0.4
0.2
DAG
TAG
Lyso PC
Lyso PE
PC
PE
PA
PG
PI
PS
Lipid Group

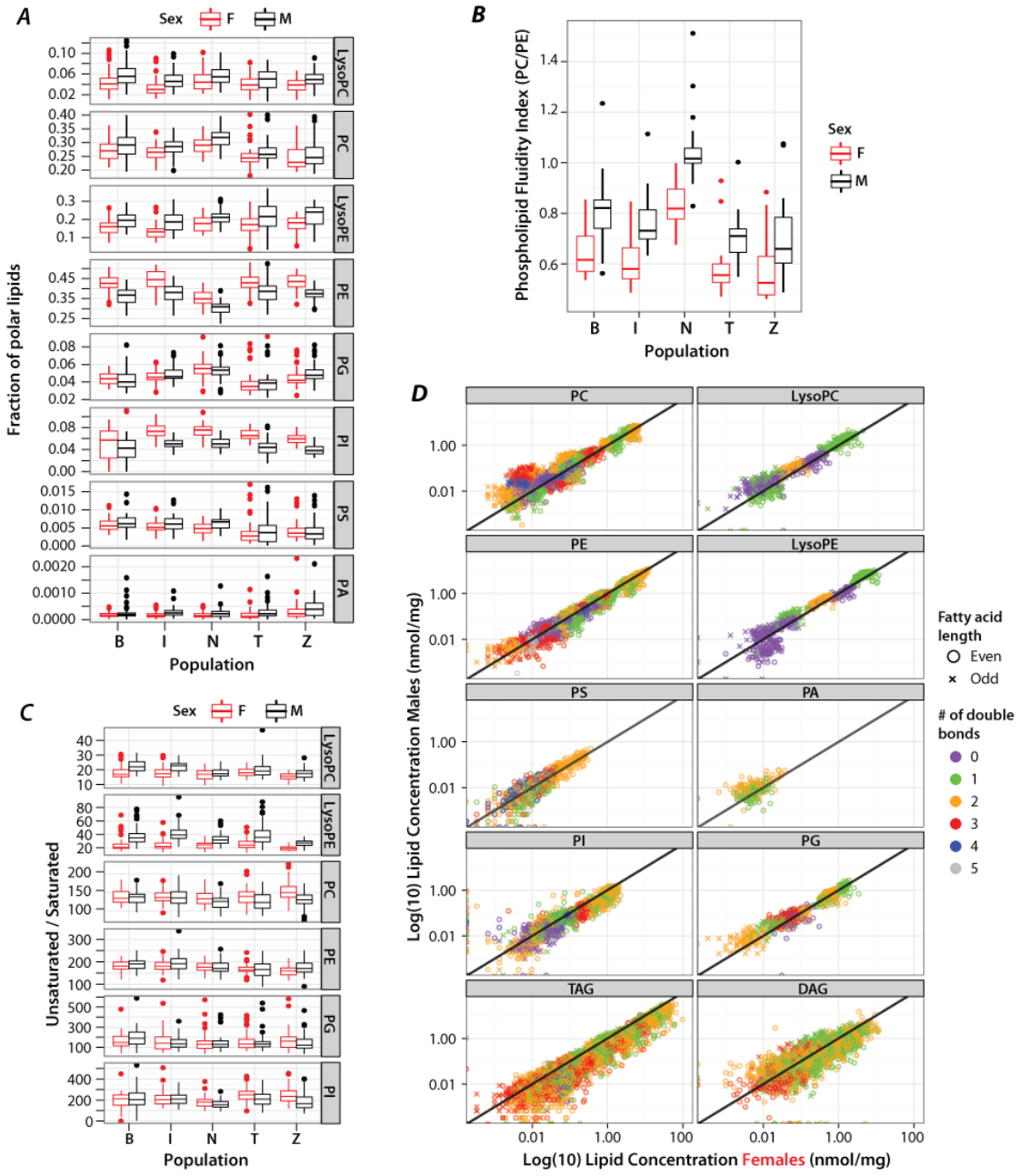
genetically. The lipid classes with the most variation in H^2 are PE and PC and below we detail potential reasons for this pattern.

3.3.2 Phospholipid ratios predict membrane fluidity

PE is the primary component of *D. melanogaster* polar lipids at all developmental stages and in all cellular membranes, with the second major membrane polar lipid class being PC (Fast, 1966; Jones et al., 1992; Hammad et al., 2011). In agreement with this we find that PCs represent ~40% and PEs ~28% of all polar lipids (Figure 3-3A). PEs are generally cone-shaped due to their smaller head-group and they cannot by themselves form a lipid bi-layer. On the other hand, PCs spontaneously organize into lipid bilayers due to their cylindrical shape. Together PEs and PCs are the two main structural components of biological membranes. Membranes that contain a low proportion of PE are said to be rigid and are at high risk to transition to the gel phase in cold-environments. Adjusting the amount of PE and PC in the inner and outer membrane leaflets is a dynamic and facile way for organisms to adjust their membrane fluidity and prevent this transition (Hazel and Williams, 1990). An increase of PE will slightly destabilize the membrane and thus maintain it in the fluid phase. As a consequence, cold-acclimated ectotherms show a lower PC/PE ratio than their warm-acclimated relatives (Hazel and Williams, 1990). Additionally, a downward shift in environmental temperature causes a drop in the PC/PE ratio in several species (Hazel and Landrey, 1988; Pruitt, 1988). Here, we detect a high PC/PE ratio in the Netherlands population, indicating rigid membranes as opposed to fluid

Figure 3-3: Distribution of lipids and membrane fluidity

A: Fraction of polar lipid classes of total polar lipid in the 5 populations and separated by sex. B: Membrane fluidity index as determined by the PC/PE ratio for all populations and sexes. C: Ratio of unsaturated over saturated lipids grouped by lipid class, population and sex. A lipid class was only included if it contained saturated lipids. D: Log-transformed concentration of all lipid species in all lipid classes in males vs females. Samples are color-coded by the number of double bonds and the FA length is categorized into odd (x) and even (circle).


A
Sex F M
Fraction of polar lipids
LysoPC
PC
LysoPE
PE
PG
PI
PS
PA
B I N T Z
Population
B
Phospholipid Fluidity Index (PC/PE)
Sex
F
M
B I N T Z
Population
C
Sex F M
Unsaturated / Saturated
LysoPC
LysoPE
PC
PE
PG
PI
B I N T Z
Population
D
PC
LysoPC
PE
LysoPE
PS
PA
PI
PG
TAG
DAG
Log(10) Lipid Concentration Males (nmol/mg)
Log(10) Lipid Concentration Females (nmol/mg)
Fatty acid length
Even
Odd
# of double bonds
0
1
2
3
4
5

membranes in all other populations (Figure 3-3B). Males also show a slightly higher ratio (more rigid membranes) compared to their female counterparts in all populations (Figure 3-3B).

Decreasing lipid saturation as summarized in the ratio of unsaturated (US) over saturated (S) lipids is an additional measure of decreased membrane fluidity (Hazel, 1995; Cossins et al., 2002). When calculating this US/S ratio for all lipids, we excluded all neutral lipids as well as PAs and PSs because we did not measure completely saturated lipids in these classes. Lysophilic lipids have the highest degree of saturation across all populations with Zimbabwe flies showing the least unsaturation observed (Figure 3-3C) due to a loss of species containing a single double bond (not shown). However, striking population differences are not observed across the two major membrane components PE and PC (Figure 3-3C) although the variation is statistically significant (Table 3-1). While lysoPEs also comprise a substantial fraction of membrane polar lipids (~19%), it is not clear how much variation in saturation of lysoPEs impacts overall membrane fluidity.

While the US to S ratio is a good simplification of the data, it does not take into account the degree of unsaturation which, clearly does have an effect on membrane fluidity (Hazel, 1995; Cossins et al., 2002). Thus we contrasted the concentrations of each lipid class in males and females grouped by the number of double bonds contained, as well as the length of the FA chains (Figure 3-3D). Generally, males tend to have more highly unsaturated lipids than females for all polar lipids except for PEs and PIs, which is not consistent with a prior study

Table 3-1: *P*-values for differences in lipid concentration or US/S ratio across population or sex for all lipid classes as calculated by ANOVA.

	Lipid Concentration		Unsaturated/Saturated ratio	
Lipid	**Population**	**Sex**	**Population**	**Sex**
PC	0.4599	8.98×10^{-4}	3.01×10^{-4}	2.34×10^{-7}
PE	$<2.2 \times 10^{-16}$	2.35×10^{-12}	2.34×10^{-7}	7.90×10^{-3}
PC/PE	$<2.2 \times 10^{-16}$	$<2.2 \times 10^{-16}$	NA	NA
PS	2.17×10^{-10}	2.21×10^{-4}	NA	NA
PI	1.87×10^{-07}	$<2.2 \times 10^{-16}$	1.02×10^{-8}	8.67×10^{-4}
PG	3.41×10^{-08}	0.530	0.058	0.69
PA	1.62×10^{-13}	3.93×10^{-08}	NA	NA
lysoPC	2.51×10^{-4}	2.61×10^{-08}	2.81×10^{-16}	5.2×10^{-16}
lysoPE	3.20×10^{-12}	2.66×10^{-11}	2.09×10^{-13}	$<2.2 \times 10^{-16}$
DAG-16.1	$<2.2 \times 10^{-16}$	7.34×10^{-10}	NA	NA
DAG-18.1	2.20×10^{-16}	2.20×10^{-16}	NA	NA
TAG-16.1	9.80×10^{-3}	2.20×10^{-16}	NA	NA
TAG-18.1	2.40×10^{-05}	2.20×10^{-16}	NA	NA

(Parisi et al., 2011). Notice, however, that our result primarily pertains to an increased US to S ratio only in lysoPCs and lysoPEs, which also show the strongest difference across the sexes (Table 3-1). In the context of membrane fluidity this is in contrast to the increased PC/PE ratio in males (Figure 3-3B) and thus could serve as a compensation mechanism to maintain membrane fluidity at levels similar to females.

There is little known about the relevance of the specific double bond configuration for each polar lipid class. Our data suggest that highly desaturated lipids are more strongly heritable compared to saturated lipids (Figure 3-4A), but more research needs to be done to elucidate the significance of this connection. However, we do know that saturated lipids make membranes more rigid and reciprocally unsaturated lipids allow for membrane fluidity and plasticity. Males have more unsaturated fatty acids overall than females, but females have a lower PC/PE ratio. These two patterns have opposite effects on membrane fluidity, potentially keeping each other in the balance. In line with this observation there is no report in the literature of any large difference in membrane fluidity of adult males vs. females.

Overall, our findings predict that Netherlands flies have more rigid membranes than any other population tested. One physiological consequence that we might predict from this observation is that Netherlands flies ought to have elevated ethanol tolerance. We decided to quantify this directly.

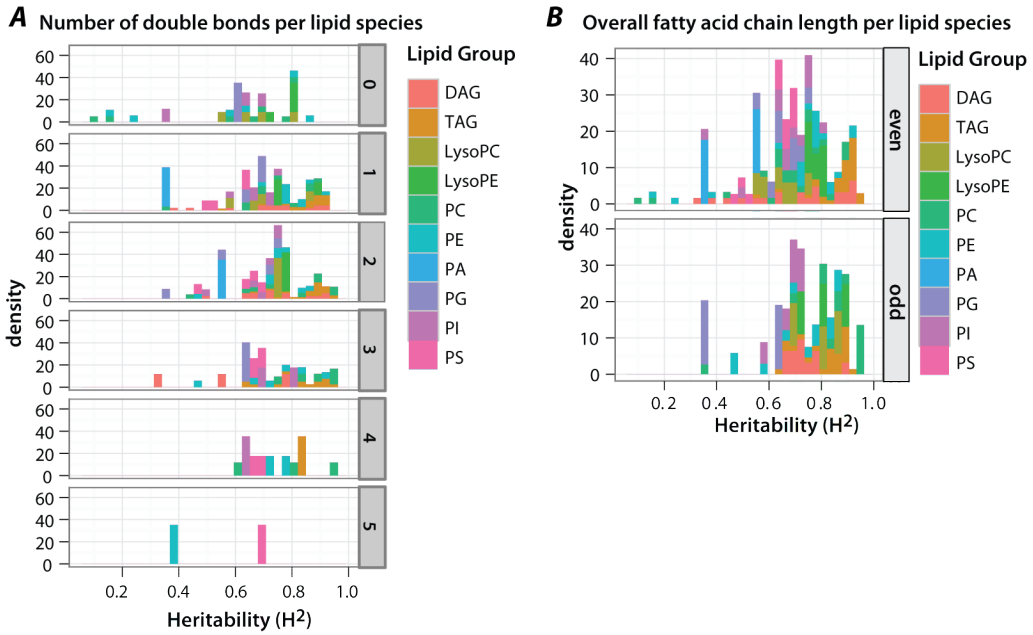


Figure 3-4: Heritabilities vary by carbon-chain length and double bond conformation

Heritability of lipid species grouped by double bond count (A) as well as odd and even chain FAs (B) per lipid class. Lipid classes are color-coded.

3.3.3 Netherlands flies show increased ethanol tolerance

Previously, Montooth et al. (2006) reported that ethanol tolerance is correlated to environmental features such as rearing altitude and temperature and thus membrane fluidity in flies. To test our hypothesis we compared ethanol resistance in Netherlands to Tasmanian flies. Four lines for each population were randomly selected and reared as before, but instead of extracting lipids we placed flies in vials with sucrose and a graded concentration series of ethanol to determine the lethal dose (LD_{50}) after 48 hours. Using probit regression we found that indeed Netherlands flies reach their LD_{50} at around 13% ethanol, while Tasmanian flies show 50% lethality already at 10% ethanol (Figure 3-5). This difference between the two populations is significant with an associated *P*-value of 6.31×10^{-11} supporting our hypothesis derived from our membrane fluidity measure that the Netherlands flies have a higher ethanol tolerance. There was no statistically significant difference in ethanol tolerance of males and females in this assay.

3.3.4 Minor phospholipids with a major impact on cellular metabolism

In addition to the prominent PE and PC lipids PG, PI, and PS lipid classes also show significant variation across populations (Figure 3-3A and Table 3-1). These phospholipids are important biological regulators and signaling components. PGs themselves only have a few known functions, such as coating lung aveoli in mammals (Okano and Akino, 1979) and forming cardiolipins, which is the integral component of mitochondrial membranes anchoring cytochrome C

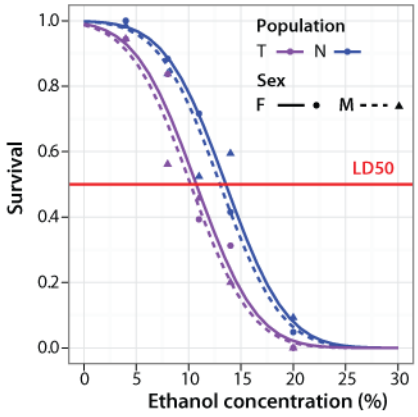


Figure 3-5: Netherland flies have higher ethanol tolerance

Ethanol tolerance measured in duplicates of 4 lines across male and female flies of each Dutch and Tasmanian origin. Red line indicates a survival rate of 0.5.

oxidase (complex IV) (Claypool and Koehler, 2012). Unfortunately, in our ESI-MS/MS short chained FAs (~10:0) in PGs and TAGs elicit very similar profiles in terms of mass of intact ions and neutral fragments and hence they are hard to identify with confidence. Hence, we decided not to further investigate differences in this lipid class.

A well-known lipid with broad biological impact is PI. Phosphorylation of the inositol ring at the 3rd hydroxyl group by PI-3 kinase is the most common alteration of PIs starting the PI3K/Akt/mTor pathway (Skwarek and Boulianne, 2009). This is crucial for many different cellular functions including, cell growth, proliferation, motility and survival, and its mis-regulation is strongly associated with cancer (Skwarek and Boulianne, 2009). While a lot is known about the enzymes involved and the downstream players after the initial phosphorylation, it is not known whether the length of the carbon chains in the respective fatty acids influence the efficacy of signaling. In our dataset 36:2 is the most abundant PI and variability across populations within PI is small compared to other lipids, albeit significant. However, it is striking that among the Beijing lines there is a wide variability in PI content (Figure 3-3A). It originates from significant line variance (*P*-value = 1.37×10^{-7}), however some lines do show broad but statistically non-significant replicate variance (*P*-value = 0.55). Among all other populations only Tasmanian flies show a marginally significant line variation (*P*-value = 0.041). Overall, the Beijing population would be a good set of flies to study the influence of lipid configuration on PI3K.

Lastly, apoptosis research usually focuses on the enzymes involved and the players involved in apoptotic cell clearance. In cells that undergo apoptosis PSs are flipped from the inner membranes leaflet to the outer so that it can be recognized by annexin V thereby recruiting macrophages for cell clearance (Vance, 2003). While many details are known about the individual proteins, it is not known whether only certain length PSs can be flipped or recognized by annexin V. It is also not clear whether absolute amounts of PIs can serve as an indicator for apoptosis itself or at least the potential for it. All the populations have significant inter-line variation in PIs. More striking is that Tasmanian and Zimbabwe flies have a significantly lower PS levels than the other populations (Figure 3-3A) and this could indicate that these two populations have less apoptotic cells to clear. Many factors influence the regulation of apoptosis but one is bacteria-induced stress. Both pathogenic and symbiotic bacteria can cause elevation of reactive oxygen species, or increased DNA breaks, resulting in apoptosis (Punj and Chakrabarty, 2003). However, symbiotic bacteria also supply unique lipid resources to the host.

3.3.5 Ratios of odd and even chain fatty acids are environmentally influenced

Bacteria have the ability to synthesize odd-chain lipids and through ingesting bacteria and their products in their food animals can acquire this class of lipids. Symbiotic interactions between bacteria and their hosts are highly heritable and this is potentially reflected in the strong heritability of odd-chained lipids (Figure

3-4B) that is statistically significantly higher than for even-chained lipids (*P*-value: $2.2x10^{-16}$).

From the ESI-MS/MS results we can extract the total number of carbon atoms within each lipid. In lyophilic phospholipids there is only one FA and so FA length is unambiguously defined. In phospholipids there are two FA chains, hence lipids with an odd total number of FA carbon atoms contain one odd-chain and one even-chain FA, while lipids with an even number of total carbon atoms could contain either two even-chain FAs or two odd-chain FAs. We know that odd-chain lipids are the minority because only bacteria can synthesize them and so lipids with an even total are far more likely to have two even-chained FAs. Synthesis and lengthening of FAs relies on the sequential addition of two carbon atoms each time, hence there will also be no even to odd conversion. Here, we have calculated the fraction of odd-carbon lipids for each phospholipid class. PA and PS lipids are excluded as no odd-carbon lipids were measured in these two classes.

While there is no significant variation of odd lipids across populations, it is striking that males have a significantly greater fraction of odd-carbon lipids than females for all lipid classes (Figure 3-6A). This is due to the increased concentration of mostly even neutral lipids in females, and also to the increased concentration of odd lipids in males across most lipid classes (Figure 3-3D). Both males and females were reared in the same environment and on the same food, and so differences in metabolism must account for this. Female flies have a larger

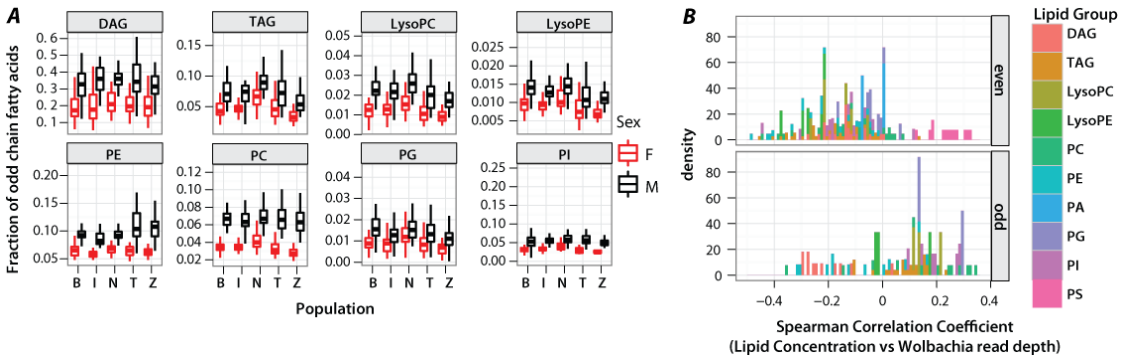


Figure 3-6: Odd-chain lipid concentration correlates with *Wolbachia* abundance

A: Fraction of odd-chained FAs estimated from the total carbon count per lipid species. Only lipid classes containing odd length FAs are included. B: Histogram of lipid species concentration correlation with normalized *Wolbachia* genome read depth (spearman) for each fly line where *Wolbachia* was detected (*n*=65). Lipid species are color-coded by lipid group and split by FA length categorized into odd and even.

fat body and harbor more storage lipids to supply the reproductive organs and eggs. However, it is not clear why males would have more odd-chained lipids.

Generally odd-chained lipids originate from external sources such as bacteria, so we next correlated lipid concentration with *Wolbachia* abundance. We approximated this abundance by using whole-genome sequencing data, using the depth of sequencing reads that aligned to the Wolbachia genome as a proxy for Wolbachia bacterial density. For this analysis we excluded the lines that did not contain *Wolbachia*. The genome sequencing data were available for females only. Odd lipids show a positive correlation with *Wolbachia* whereas abundance of even lipids showed no or negative correlations (Figure 3-6B, *P*-value = 7.47×10^{-13}), suggesting that *Wolbachia* does indeed promote the accumulation of odd lipids. In addition, one class of lipids that only contains even chained FAs is also correlated with *Wolbachia* abundance, PS (Figure 4B), the class responsible for signaling to clear apoptotic cells. Notably, without PS the difference in correlations between odd and even-chained lipids becomes stronger (*P*-value = 2.20×10^{-16}) whereas the distribution of PS does show a slight location shift towards stronger correlations (*P*-value = 8.17×10^{-3}). This suggests that while having a new class of lipids available *Wolbachia* presence also presents a significant amount of stress to the organism, increasing its apoptosis, which is in line with a recent study in *D. simulans* (Brennan et al., 2012). Furthermore, since all lengths of PS are positively correlated, our data suggests that they all may play a role in induction of apoptosis.

3.3.6 Correlation Structures and sex dichotomy

Our dataset confirms that females have a large quantity of storage lipids (TAGs see Figure 3-3D). While all neutral lipid classes show statistically significant population variation (Table 3-1) the general pattern across the classes is constant. Within neutral lipids males, in general, have a higher fraction of 16:1 containing DAGs at the expense of 18:1 containing TAGs (Figure 3-9A). Across all polar and neutral lipids we find a total of 72 lipid species that show clear sex-differences in their concentration. 43 of those show strong signatures of high vs. low concentration in females vs. males and vice versa. Notably, some lipids have a reversed pattern in different populations (PIs, Figure 3-7A). Also high concentrations of, for example DAG 37:2 (16:1), are strongly male-biased, whereas other DAG species (DAG 34:1 (18:1)) are female biased, as is the entire class of DADs (Figure 3-7B). In addition to analyzing the individual lipid classes we also determined the clusters of intercorrelated lipid classes that form in each population. To compare sexes we removed the population contribution and grouped lipids using modulated modularity clustering (Ayroles et al., 2009; Stone and Ayroles, 2009). Strikingly, in the males there were 72 weakly supported clusters, each containing 3 lipids on average. Hence, there is no significant grouping of lipids observed, not even according to the lipid head groups (Figure 3-8A). On the other hand females form 13 clusters mostly determined by lipid class and odd or even chain FAs (Figure 3-8B). Of particular interest is cluster 10 because it contains only odd-chain FAs, but it includes all of the distinct polar

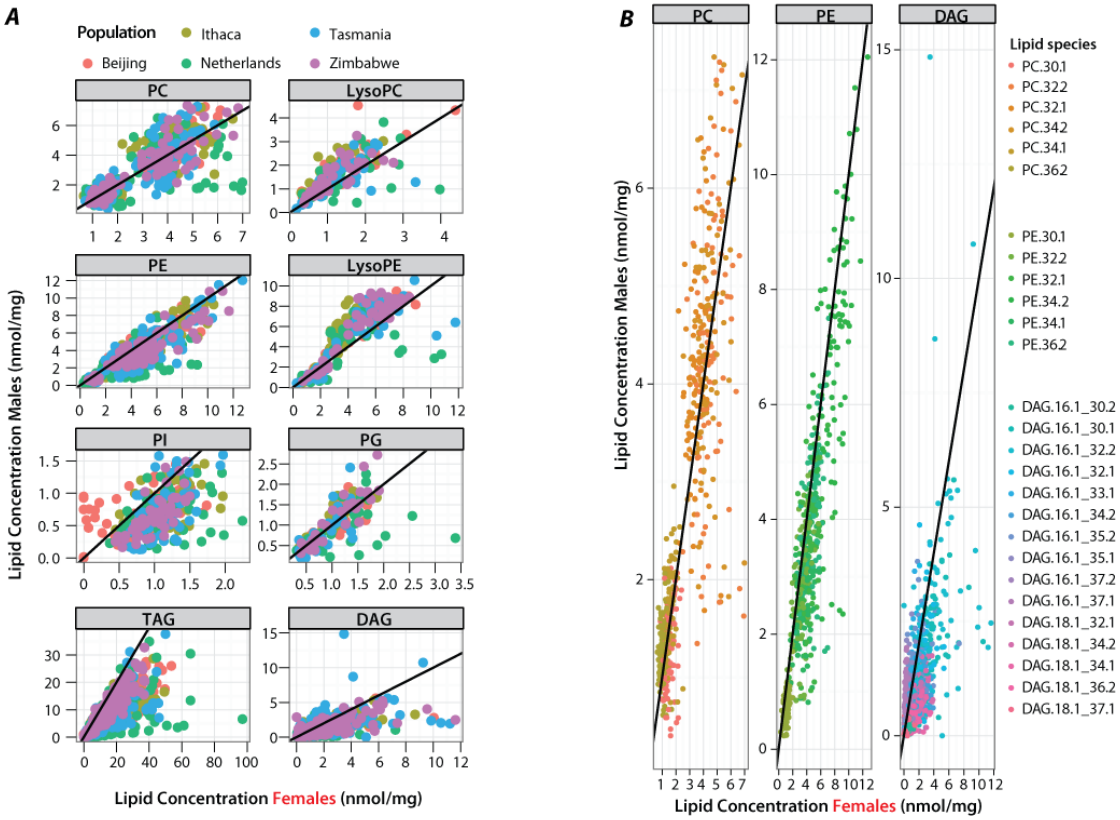


Figure 3-7: Striking differences in lipid concentrations across the sexes

A: Concentration of lipid species that show sex differences (>1) in all lipid classes in males vs females. Samples are color-coded by population. B: Subset of A with lipid species-specific behavior. Samples are color-coded by lipid species.

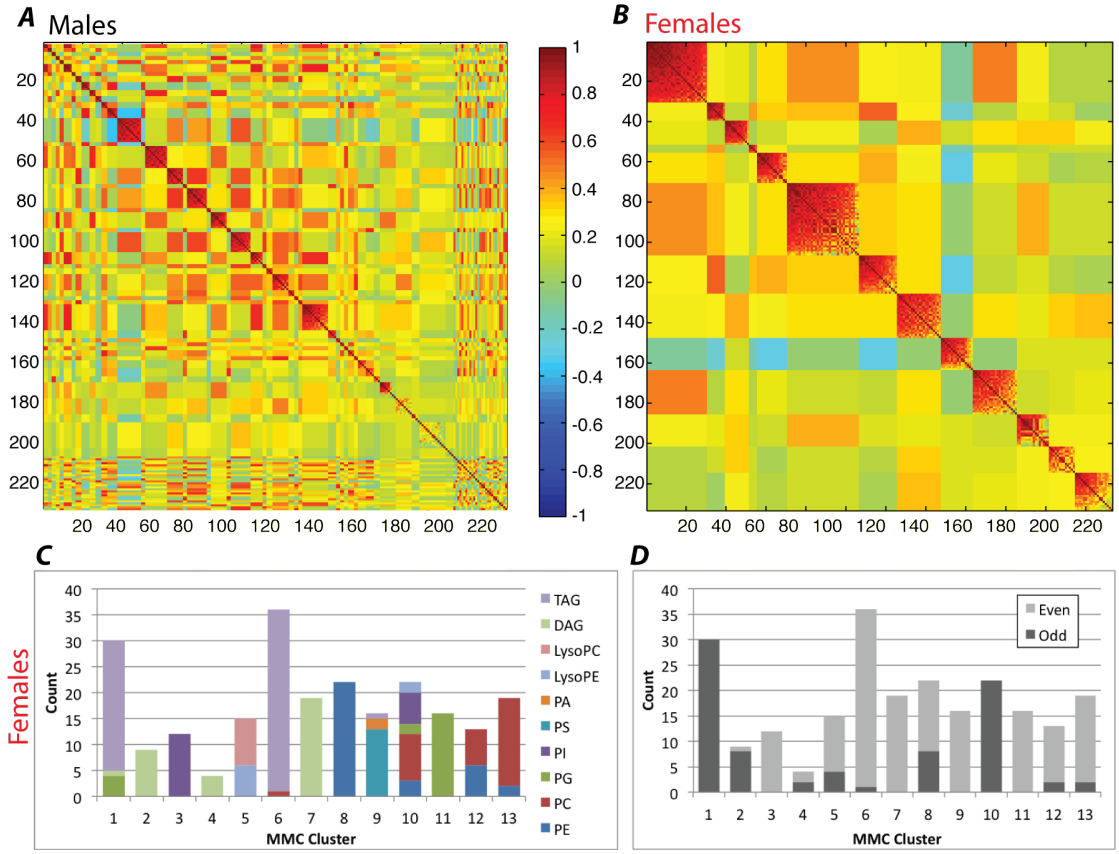


Figure 3-8: MMC clustering reveals distinct clusters only in female lipid data

A and B: MMC matrix for male (A) and female (B) lipid concentration with any population effect removed. C: Lipid class distribution across identified clusters in females (from B). D: Odd and even chain FA distribution across the same clusters.

lipid classes. This composition is unique among the clusters we find (Figure 3-8C,D), but the biological significance is not clear. This cluster possibly derives from the fact that odd-chain lipids have to be drawn from a limited pool that is acquired through food or symbiotic bacteria. Other correlations confirm the top-level structure presented above, such as the negative correlation of DAGs with its derivative lipids (compare Figure 3-2B to cluster 9 of Figure 3-8B).

We next carried out discriminant analysis of principle components (DAPC) using our dataset for males and females without the population component (Figure 3-9B). Retaining 20 principal components (PCos) we see a clear separation between the sexes. The set of lipids with a component loading of greater than 0.2 contains several TAGs and DAGs as expected, but also various signaling and membrane lipids (Figure 3-10A).

3.3.7 Discriminant analysis of principal components reveals significant differences in population structure between males and females

The sex disparities prompted us to investigate population structure in males and females separately using DAPC. To achieve a clear separation of populations we retained 30 PCos in either sex for the discriminate analysis. In both males and females the components explain slightly over 80% of the cumulative variance in the sample (Figure 3-9C,D). There are prominent differences between the sexes in their population distribution. The males show one clear outlier, the Beijing population, and all other populations are equally spaced from one another (Figure 3-9C,D). In line with a clear separation, grouping populations according

Figure 3-9: DAPC analysis reveals clear population separation in either males or females

A: Fraction of neutral lipids across population and sex. Value in brackets identifies one FA of a total 2 for DAGs and a total 3 for TAGs. B: DAPC analysis for all flies investigating the sex separation. Inlay shows the total variance explained by the PCos. Retained PCos are colored black. C and D: DAPC analysis for males (C) and females (D) separating the 5 populations. Colored crosses and their connections show the minimum-spanning tree based on the population distances. Df = Discriminate function. Inlay as in B. E: Percent of individuals that are assigned to their biological population using the DAPC of C and D. F: Leave-one-out approach testing the robustness of the population assignment from the DAPC analysis in C and D. G: eigenvectors from the male DAPC (C) are applied to the female lipid dataset and it was determined how well this allows for population separation by measuring the re-assignment of individuals to their biological population. H: Same as F but female derived Eigenvectors are applied to the male dataset.

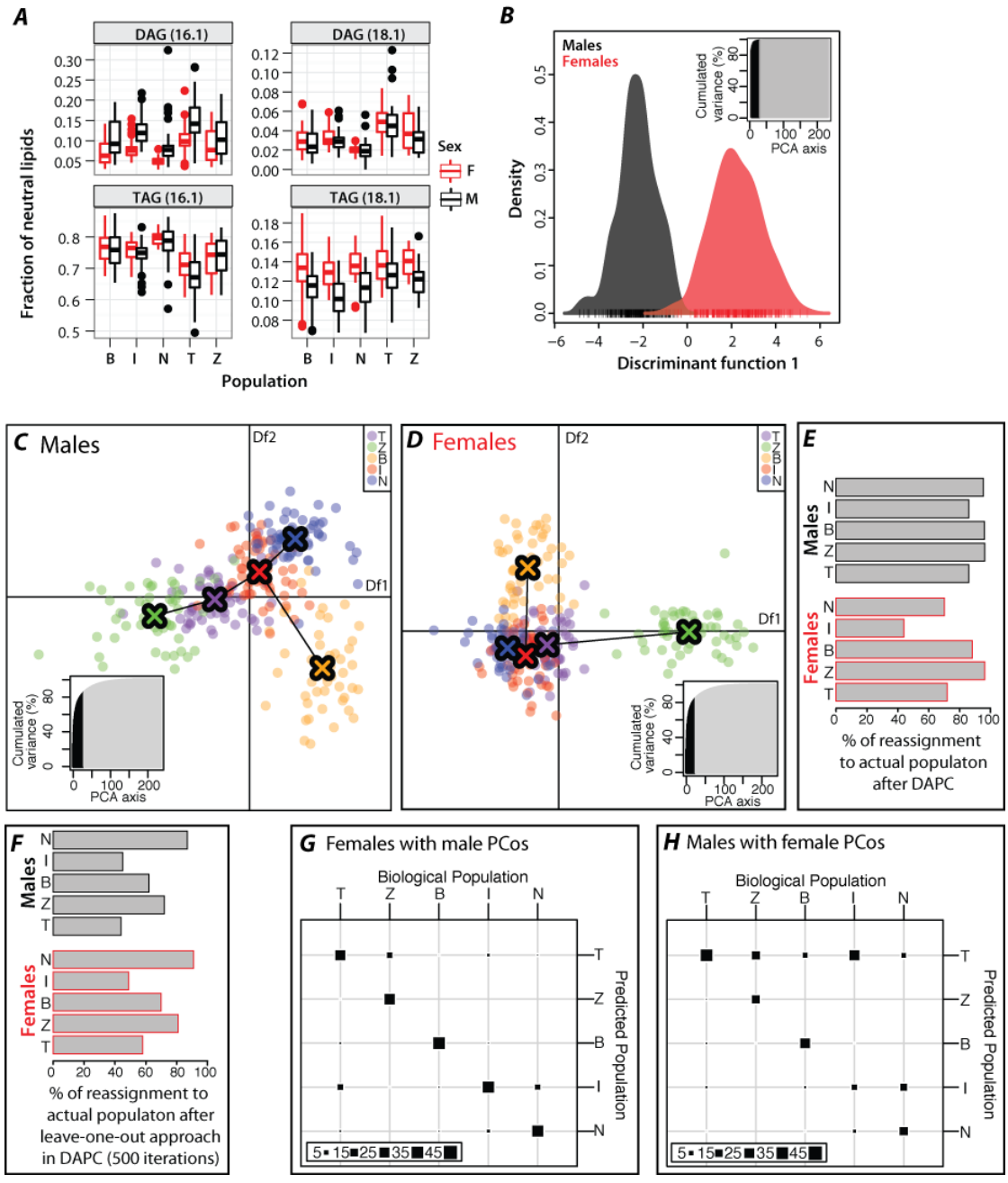
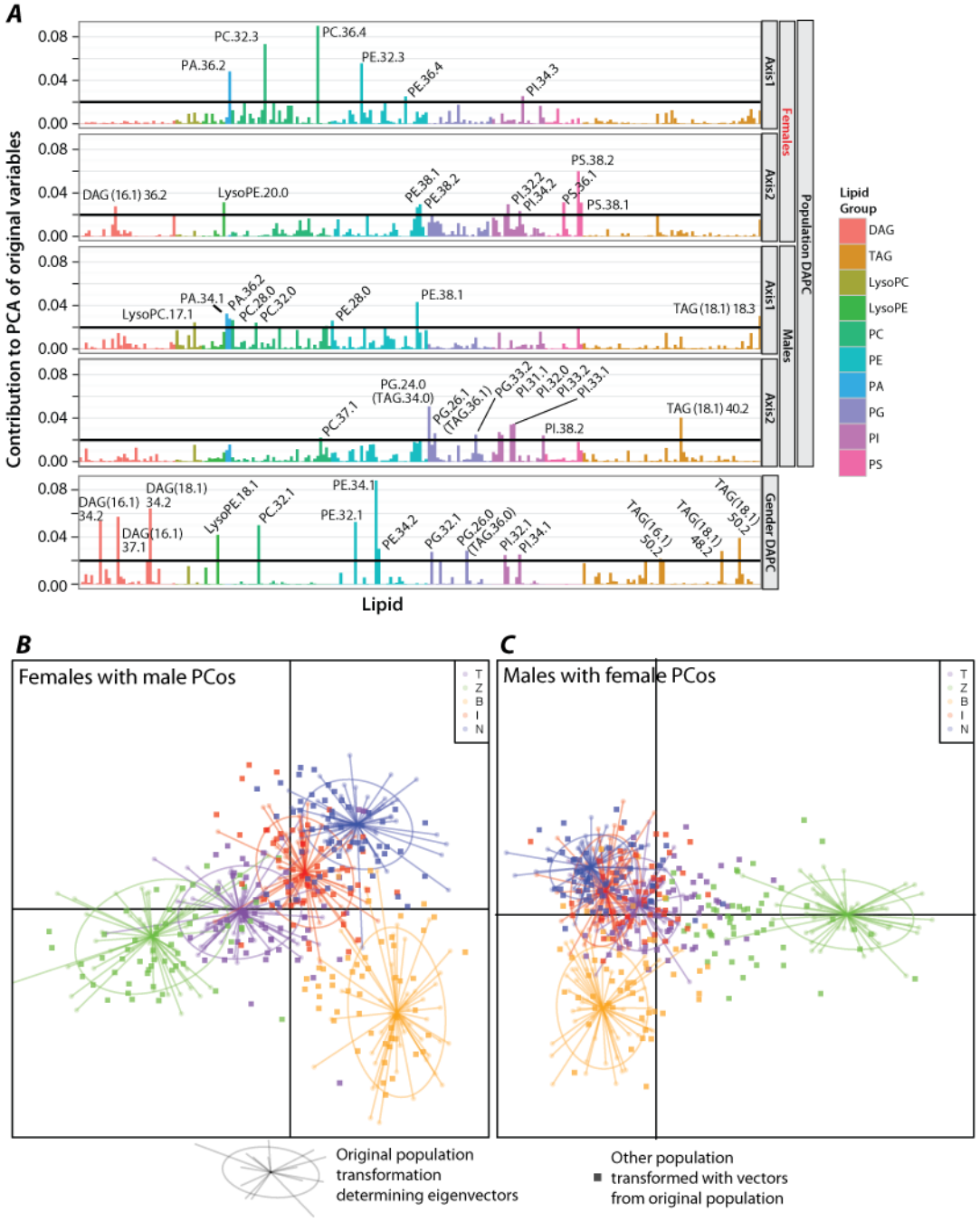

A
DAG (16.1)
DAG (18.1)
TAG (16.1)
TAG (18.1)
Fraction of neutral lipids
Population
B T I N T Z
Sex
F
M
B
Males
Females
Density
Discriminant function 1
Cumulated variance (%)
PCA axis
C Males
D Females
Df1
Df2
E
% of reassignment to actual populaton after DAPC
F
% of reassignment to actual populaton after leave-one-out approach in DAPC (500 iterations)
G Females with male PCos
H Males with female PCos
Biological Population
Predicted Population
5 15 25 35 45

Figure 3-10: Lipids that load the DAPC are strikingly different across sexes

A: Loading of original variables for DAPC from Figure 3-9B-D. For population DAPCs both discriminate functions are shown (Axis 1 and 2). Lipids are colored by lipid class and individual lipid species with a component loading greater than 0.2 (black line) have been annotated. B and C: Scatter plots of cross-sex DAPC Eigenvector transformation as summarized in Figure 3-9G,H.


A
Contribution to PCA of original variables
Lipid
Axis1
Axis2
Females
Males
Population DAPC
Gender DAPC
Lipid Group
DAG
TAG
LysoPC
LysoPE
PC
PE
PA
PG
PI
PS
PC.32.3
PC.36.4
PA.36.2
PE.32.3
PE.36.4
PI.34.3
DAG (16.1) 36.2
LysoPE.20.0
PE.38.1
PE.38.2
PI.32.2
PI.34.2
PS.36.1
PS.38.2
PS.38.1
LysoPC.17.1
PA.34.1
PA.36.2
PC.28.0
PC.32.0
PE.28.0
PE.38.1
TAG (18.1) 18.3
PC.37.1
PG.24.0
(TAG.34.0)
PG.26.1
(TAG.36.1)
PG.33.2
PI.31.1
PI.32.0
PI.33.2
PI.33.1
PI.38.2
TAG (18.1) 40.2
DAG(16.1) 34.2
DAG(18.1) 34.2
DAG(16.1) 37.1
LysoPE.18.1
PC.32.1
PE.34.1
PE.32.1
PE.34.2
PG.32.1
PG.26.0
(TAG.36.0)
PI.32.1
PI.34.1
TAG(16.1) 50.2
TAG(18.1) 48.2
TAG(18.1) 50.2
B
Females with male PCos
C
Males with female PCos
T
Z
B
I
N
Original population
transformation
determining eigenvectors
Other population
transformed with vectors
from original population

to DAPC matches the original population assignment for at least 80% of lines in each population (Figure 3-9E). Strikingly, the females show two outlier populations; Zimbabwe and Beijing. However, Ithaca, Tasmanian and Dutch flies map very close together, although still in the same order as in males, resulting in ill-defined populations boundaries between the three of them. The lack of population boundaries is mostly affecting the population of Ithaca flies, which is only reassigned in 45% of the cases (Figure 3-9E). We also tested the robustness of the group assignment in each sex determined by DAPC. We computed eigenvector and eigenvalues leaving out one line per population and determined the group the additional line fell into given the computed PCos. The success of the leave-one-out approach is dependent both on the variation within a population (cluster strength) and among population (cluster distance). In the Netherlands males, for example, the leave-one-out misclassification error rate was less than 10%, meaning that more than 90% of the time, one could successfully identify a fly from this population based only on its lipids (Figure 3-9F). For each population, females showed a slightly better recovery, indicating that although populations are more overlapping, they are more tightly clustered and thus removing one line will less affect the boundaries defining that group (Figure 3-9F).

To understand the underlying differences between male and female PCos we looked at the contribution of the individual lipids to the population discrimination. As expected by our previous analysis there is little overlap

between the top contributing lipids between the sexes and neither set of female or male loading lipids intersects with the set discriminating sex except for PI 34:2 (Figure 3-10A).

While it is expected that lipid biology is drastically different in males and females, we still expected some similarities across populations. Thus we hypothesized that maybe some of the components could also differentiate between the populations in the other sex, but were only secondary to the ones driving the population differences. Consequently we asked if male PCos could recover female population structure and vice versa and the results are summarized in Figure 3-9G,H and Figure 3-10B,C. Interestingly, females fall into their biological populations when rotated with male PCos. However, males did not group well with female PCos, resulting in the misclassification of all populations, especially assigning Ithaca and Zimbabwe lines to the Tasmanian population (Figure 3-9G). These results suggest the presence of an additional component of lipid biology present in females, whereas the basic lipid biochemistry surfacing in the male DAPC is shared across sexes. This is not unexpected due to the presence of additional fat body and related functions in female reproduction.

3.4 Discussion

We are just beginning to understand the interactions of environment and genotype and its effect on lipid abundance and membrane phospholipid composition. This is the most comprehensive analysis of lipidomics in natural lines of *D. melanogaster* to date. The lipid species found in our profile are mostly

in agreement with a prior report of Hammad et al. (2011) using a new LC-MRM approach (liquid chromatography multiple reaction monitoring). Across all lipids C_{16} and C_{18} FAs dominate (Overgaard et al., 2008; Hammad et al., 2011) which in our study translates to dominance of a total of 34 carbon atoms across the FAs within a lipid species. However, they do not find phospholipids longer than C_{18}, but we detect small amounts of what have to be C_{20} PC and PE and minor amounts in PI and PS. Additionally we find TAGs with C_{20} and even C_{22} for DAGs. Parisi et al., (2011) also detects C_{20} but not C_{22} FAs. All lipid species with very long FAs have small concentrations and other prior studies indicate that these long-chain FAs exist in *D. melanogaster* (Koštál et al., 2011; Parisi et al., 2011). Given that Hammad and colleagues used a different detection method, this finding indicates that the LC-MRM approach is biased towards short-chain FAs and that mass spectrometry could give a more complete profile.

Using this rich dataset we can for the first time assess the heritability of the storage and signaling lipid components in metabolism. Overall, the heritability is very high for all classes, but highest for TAGs, the storage lipid component. Notably, PCs and PEs have both strongly and weakly heritable species. The broad distribution of PC and PE heritability could be the result of adaptation to the new environment since heritability can significantly change upon environmental stress, such as a sudden temperature and climate change during the transition to the laboratory environment (Charmantier and Garant, 2005). On the other hand,

this pattern could have stayed constant across this transition and simply reflect a high membrane plasticity that compensates for environmental fluctuations.

In addition to their heritability we observe the most striking population difference in the PC/PE ratio, where Dutch flies show a clear signature of highly structured membranes as opposed to fluid membranes for all other populations. Phospholipids are necessary for proper signaling but more importantly form the cell membranes, which are the physical barriers between the extra- and intra-cellular compartments (Cossins and Raynard, 1987; Hazel, 1995). PCs and PEs are the main structural membrane components and they also function as gatekeepers to allow for selective transport across membranes enabling proteins and nutrients to enter the cell (Fast, 1966; Jones et al., 1992; Hammad et al., 2011). To confirm that Dutch flies have less fluid membranes several methods were available. Prior plasticity in membrane fluidity confers an increased survival rate upon cold-tolerance and cold-shock assays, but the detailed effects of cold-tolerance and cold-shock on membrane lipid composition are vastly different (Tomcala et al., 2006; Overgaard et al., 2008; Koštál et al., 2011). However, there is no standardized test for cold-shock or cold-tolerance and each proposed test has significant set-up and experimental variation, leading to conflicting results from various studies (Sinclair and Roberts, 2005; Koštál et al., 2011). In addition starvation influences the outcomes since cold-treatment and recovery are routinely carried out in empty vials. Thus we used the alternative measure of ethanol resistance to verify our conclusions. Overall, membrane

fluidity is influenced by temperature but also dietary constraints. Dietary ethanol will integrate into the membrane and ethanol degradation alters the PC and PE regulation dramatically increasing its fluidity (Rubin and Rottenberg, 1982; Sun and Sun, 1985; Taraschi and Rubin, 1985; Hoek and Rubin, 1990; Gustavsson, 1995; Baker and Kramer, 1999; Shukla et al., 2001). Thus more individuals with more rigid membranes show a higher ethanol tolerance. This assay is easily standardized and the sucrose-ethanol mix provides a basic diet. Using the ethanol resistance measure we confirmed that Dutch flies have more rigid membranes, because they are more tolerant to ethanol integrating into their membranes.

Still the question remains how these differences have developed. There are two possibilities a) they can be inherited and remain constant across the transition to the laboratory environment, and b) they may have developed upon the transition of the original environment to the laboratory, and are thus an example of homeoviscous adaptation. Cold-tolerance results in an overall shortening of FAs reflected in an increased ratio of 16:1 FAs (Ohtsu et al., 1998; Overgaard et al., 2005) and this effect was also seen in *D. melanogaster* populations along a south-north gradient (Ohtsu et al., 1998), supporting the argument that these are long-term changes. Unfortunately we cannot distinguish between the individual FAs in a phospholipid to analyze if we see a decreased amount of 16:1 FAs in Netherlands flies, but the total number of carbon atoms in a phospholipid is not significantly higher in Netherlands flies compared to the other populations. However, the effects of long-term cold-acclimation go well beyond membrane

fluidity affecting basic metabolic processes such as cell cycle regulation, the production of anti-freeze proteins and adaptation of protein folding (Zachariassen and Kristiansen, 2000; Duman, 2001; Koštál, 2006; Rinehart et al., 2007; Koštál and Tollarová-Borovanská, 2009). In fact these changes can alleviate the requirement for membrane fluidity adjustment. It is possible, that our Netherlands population is the only one with all the necessary alterations to cope with constantly low temperatures without showing strong membrane fluidity. On the other hand Beijing, Ithaca, Zimbabwe and Tasmanian flies adjusted by increasing their membrane fluidity, developing a form of cold-tolerance. In agreement with prior studies of cold-acclimation our data reveals a decrease in PC/PE ratio, but no changes in the US/S ratio (Overgaard et al., 2008; Koštál et al., 2011). At this point we cannot conclusively determine if cold-acclimation as a result of the shift to the constant laboratory temperature did affect the membrane composition. Alternatively, the difference between the Netherlands and the remaining populations is not a result of cold-acclimation, but simply reflects inherited differences. A study by Cooper et al. (2011) showed that long-term rearing of two sub-populations are two different environmental conditions did not lead to a divergence in PC/PE ratio. Along this line, the differences in membrane composition could just be long-term inherited. To identify whether the difference in PC/PE ratio is strictly inherited, due to cold-acclimation or a combination of the two, one would need to determine the exact FA changes and gather information about the environmental conditions at the original collection site.

To date there is significantly less know about the biological significance of variation in the other signaling lipid classes PG, PI and PS. Especially PG has been little studied although it potentially has a tremendous biological impact. In our dataset we could not unambiguously identify PGs versus TAGs, but considering overall levels Netherlands flies have the biggest pool of PGs (Figure 2C). Mutants in the cardiolipin biosynthesis pathway suggest that PG and cardiolipin content are positively correlated (Claypool and Koehler, 2012) and thus a bigger pool of PGs could be indicative of more mitochondria and thus reflect an up-regulation in energy and heat metabolism. This connection should be evaluated further in the future, as mitochondria are crucial to cells and overall body condition. Similarly, the influence of PI double bond conformation and FA length on PI3K induction is not known, but we find significant population variation. PIs that are required to trigger the PI3K pathway, which has been extensively studied in the context of cancer formation (Cully et al., 2006). It has to be a goal for the future to fill that gap of knowledge and start adding lipidomics to our research.

Our analysis on *Wolbachia* content is a prime example of how lipidomics can foster our understanding of cell biology and cellular interactions with the environment. Intriguingly, we find that odd lipids are significantly associated with *Wolbachia* abundance, confirming that bacteria are needed to regulate the odd-chain FA supply. This raises the question which other bacteria in *D. melanogaster* provide odd-chain FAs and if there is a dominating species. Whole-fly bacterial concentration measurements would yield a better resolution that

can be used to gain leverage at the network that likely exists. Surprisingly our data show that *Wolbachia* load is also positively correlated with PS content, notably a lipid class that only contained even-chained FAs in our dataset. PSs trigger macrophage clearance of apoptotic cells by being presented to the extra-cellular space where annexin V can recognize them. The most straightforward interpretation for the correlation of PS lipids with *Wolbachia* content is that presence of this endosymbiont also causes cellular stress that ultimately results into an increased need to clear apoptotic cells. A similar connection has been made in *D. simulans* where *Wolbachia* presence in spermatocytes was associated with an increased in DNA breaks and apoptosis (Brennan et al., 2012). Notably, this association of all PS species with *Wolbachia* abundance also suggests that all members of this class can signal apoptosis, disregarding the double bonds status or FA length, albeit with different efficiencies. Moreover it proposes that PS concentration could indeed be a measure of apoptosis.

Lastly, we investigated the population structure apparent in the lipid data. A prior report from Parisi et al. (2011) reported little differences across the sexes, which is clearly not the case for our dataset. It is of note that both Parisi et al. (2011) and this study use the mass spectrometry approach to quantify lipids. The discrepancies in sex differences across lipids are likely attributable to the fact that Parisi and colleagues analyzed a single laboratory strain and our dataset is based on 92 lines with natural variation. In addition to striking high/low concentration differences for individual lipid species across males and females

we find that modularity clustering leads to two drastically different correlation patterns. It is possible, that the small clusters found in males are subgroups within the larger clusters found in females. Additionally, DAPC analysis indicates that females have an additional layer of lipid biology that is absent in males. Eigenvectors derived from male lipid profiles can be used to transform the female data, leading to a very similar population separation, however the reverse is not true. This is inline with the clusters observed in the MMC analysis and indicates that the distinct structures in males and females are due to differences across all lipid classes and their interactions. On the other hand, analyzing the strongest differences between the lipid profiles in males and females only points toward species of all lipid classes with neutral (storage) lipids showing the most differences. This is consistent with females having additional fat body tissue. Overall, this demonstrates that there are more subtle differences between male and female lipid metabolism that have a big effect on the observed population variance.

There is a lot to learn about the biological significance of lipid variation. From decades of "omics" research we have seen repeatedly that little is left to chance when it comes to the molecular biology of the cell. So far research has focused on proteins and their interactions with other proteins, and with DNA and RNA. The analysis of variation in lipids lags behind, but our demonstration of the strong signatures of inter-line variability, of heritability and of modular clustering of correlations beg for a deeper functional analysis.

3.5 Materials and Methods

3.5.1 Fly populations and experimental design

We obtained lipid measures from 92 lines derived from 5 populations of *D. melanogaster*; Beijing(B), Ithaca(I), Netherlands(N), Tasmania(T), and Zimbabwe(Z).

The lines were derived from wild-caught females that were inbred for 13 generations by sib mating (previously described by Greenberg et al., 2010). Overall, they were grown for more than 15 generations at constant room temperature prior to lipid extraction. For the experimental generation all lines were grown on Cornell BLA media at 25°C. Males and Females were separated 3 days after eclosion and reared for another 3 days. Flies were frozen in sets of 10 for lipid extraction, or placed in vials for ethanol tolerance measurement.

3.5.2 Lipid Extraction

Lipids were extracted from 3 biological replicates per fly sex and line using a modified Bligh-Dyer lipid extraction (Bligh and Dyer, 1959). Briefly, flies were weighed before and after the 5-step extraction. To maintain the signal of polyunsaturated lipids 0.01% butylated hydroxytoluene (Fisher) was added to the chloroform used in each step. The lipid containing chloroform was washed with 1M KCl and water before being transferred to a screw-cap vial and dried down using nitrogen for shipping. The lipid profiles were determined by the Kansas Lipodomics Research Center using ESI-MS/MS.

3.5.3 ESI-MS/MS lipid profiling

The lipid profile data were acquired at Kansas Lipidomics Research Center (KLRC). Total lipid analysis was carried out according to Welti et al., 2002 and lipids were dissolved in chloroform for analysis. The profiles of membrane lipids were measured by an automated electrospray ionization tandem mass spectrometry method as previously described Devaiah et al. (2006) or by a slightly modified procedure described here. The samples were dissolved in 1 ml chloroform. An aliquot of 40 µl of extract in chloroform was used. Precise amounts of internal standards, obtained and quantified as previously described Welti et al. (2002), were added in the following quantities: 0.30 nmol di12:0-PC, 0.30 nmol di24:1-PC, 0.30 nmol 13:0-lysoPC, 0.30 nmol 19:0-lysoPC, 0.30 nmol di12:0-PE, 0.30 nmol di23:0-PE, 0.30 nmol 14:0-lysoPE, 0.30 nmol 18:0-lysoPE, 0.30 nmol di8:0-PG, 0.30 nmol di20:0(phytanoyl)-PG, 0.30 nmol di14:0-PA, 0.30 nmol di20:0(phytanoyl)-PA, 0.20 nmol di14:0-PS, 0.20 nmol di20:0(phytanoyl)-PS, 0.47 nmol 16:0-18:0-PI, 0.33 nmol di18:0-PI, 4.65 nmol di15:0-DAG, and 3.10 nmol tri17:1-TAG. The sample and internal standard mixture was combined with solvents, such that the ratio of chloroform/methanol/300 mM ammonium acetate in water was 300/665/35, and the final volume was 1.2 ml.

Unfractionated lipid extracts were introduced by continuous infusion into the ESI source on a triple quadrupole MS/MS (API 4000, Applied Biosystems, Foster City, CA). Samples were introduced using an autosampler (LC Mini PAL, CTC Analytics

AG, Zwingen, Switzerland) fitted with the required injection loop for the acquisition time and presented to the ESI needle at 30 μl/min.

Sequential precursor and neutral loss scans of the extracts produce a series of spectra with each spectrum revealing a set of lipid species containing a common head group fragment. Lipid species were detected with the following scans: PC and lysoPC, $[M + H]^+$ ions in positive ion mode with Precursor of 184.1 (Pre 184.1); PE and lysoPE, $[M + H]^+$ ions in positive ion mode with Neutral Loss of 141.0 (NL 141.0); PG, $[M + NH4]^+$ in positive ion mode with NL 189.0 for PG; lysoPG, $[M - H]^-$ in negative mode with Pre 152.9; PI, $[M + NH4]^+$ in positive ion mode with NL 277.0; PS, $[M + NH4]^+$ in positive ion mode with NL 185.0; PA, $[M + NH4]^+$ in positive ion mode with NL 115.0; DAG internal standards species containing 15:0, $[M + NH4]^+$ in positive ion mode with NL 259.2; TAG internal standards species containing 17:1, $[M + NH4]^+$ in positive ion mode with NL 285.2; DAG and TAG containing 16:1, $[M + NH4]^+$ in positive ion mode with NL 271.2; and DAG and TAG containing 18:1, $[M + NH4]^+$ in positive ion mode with NL 299.2. The scan speed was 50 or 100 u per sec. The collision gas pressure was set at 2 (arbitrary units). The collision energies, with nitrogen in the collision cell, were +28 V for PE, +40 V for PC, +25 V for PI, PS and PA, +20 V and PG, +20 V for DAG and TAG. Declustering potentials were +100 V for PE, PC, PA, PG, PI, and PS, and +100 V for DAG and TAG. Entrance potentials were +15 V for PE, +14 V for PC, PI, PA, PG, and PS, and +14 V for DAG and TAG. Exit potentials were +11 V for PE, +14 V for PC, PI, PA, PG, PS, and +14 V for DAG and TAG. The mass analyzers

were adjusted to a resolution of 0.7 u full width at half height. For each spectrum, 9 to 150 continuum scans were averaged in multiple channel analyzer mode. The source temperature (heated nebulizer) was 100 °C, the interface heater was on, +5.5 kV or -4.5 kV were applied to the electrospray capillary, the curtain gas was set at 20 (arbitrary units), and the two ion source gases were set at 45 (arbitrary units).

The background of each spectrum was subtracted, the data were smoothed, and peak areas integrated using a custom script and Applied Biosystems Analyst software. The lipids in each class were quantified in comparison to the two internal standards of that class. The first and typically every 11^{th} set of mass spectra were acquired on the internal standard mixture only. Peaks corresponding to the target lipids in these spectra were identified and molar amounts calculated in comparison to the internal standards on the same lipid class. To correct for chemical or instrumental noise in the samples, the molar amount of each lipid metabolite detected in the "internal standards only" spectra was subtracted from the molar amount of each metabolite calculated in each set of sample spectra. The data from each "internal standards only" set of spectra was used to correct the data from the following 10 samples. Finally, the data were corrected for the fraction of the sample analyzed and normalized to the sample "dry weights" to produce data in the units nmol/mg.

For the TAG and DAG analysis mass spectral signals in the NL 271.2 and NL 299.2 scans were normalized to the signal for 4.65 nmol di15:0-DAG in the NL 259.2

scan (for DAG species), and 3.10 nmol tri17:1-TAG (for TAG species); relative mass spectral signal/lean weight for DAG was calculated by the following formula:

$$\frac{\left(\begin{array}{c}\text{mass spectral signal for indicated DAG or TAG species}\\ \times \\ \text{internal standard amount in nmol}\end{array}\right)}{\left(\begin{array}{c}\text{mass spectral signal for the internal standard}\\ \times \\ \text{fraction of sample analyzed x lean weight}\end{array}\right)}$$

Thus, the data are presented as relative mass spectral signals, where a value of 1 for a DAG or TAG represents the same amount of signal as 1 nmol of DAG or TAG standard, respectively.

3.5.4 Ethanol Resistance

10-15 flies are placed in vials containing a Whatman disk soaked in 1ml of a 3% sucrose solution supplemented with 4,8,11,14 or 20% Ethanol sealed with parafilm. Duplicate vials were placed at 25°C for 48 hours. Live and dead flies were scored and recorded for statistical analysis.

3.5.5 Statistics

The raw data were normalized and lipids lacking data or significantly deviating from a normal distribution were excluded (n=109). Additionally, one replicate had an ~8 fold increased concentration for each lipid and was thus excluded. This resulted in one line having only two replicates, all other 183 lines have three replicates. Batch effects from lipid extraction and ESI-MS/MS profiling were

analyzed and removed using PVCA (Boedigheimer et al., 2008; Scherer, 2009). The count of *Wolbachia* sequence reads were normalized to the line's genome read depth.

To compare population variances or sex variance an ANOVA accounting for between and within subject variance was used. DAPC was carried out using the adegenet package from R (Jombart et al., 2010). Population effects were removed with REML where indicated. To determine ethanol resistance LD_{50} values were estimated and analyzed using a generalized linear model with probit regression. The goodness of fit was determined with a likelihood ratio test. Differences in histogram spread were analyzed using the Wilcoxon rank sum test with continuity correction in R.

REFERENCES

Ayroles JF, Carbone MA, Stone E a, Jordan KW, Lyman RF, Magwire MM, Rollmann SM, Duncan LH, Lawrence F, Anholt RRH, Mackay TFC. 2009. Systems genetics of complex traits in Drosophila melanogaster. *Nature Genetics* **41**: 299–307.

Baker RC, Kramer RE. 1999. Cytotoxicity of short-chain alcohols. *Annual Review of Pharmacology and Toxicology* **39**: 127–150.

Bligh EG, Dyer WJ. 1959. A rapid method of total lipid extraction and purification. *Canadian Journal of Biochemistry and Physiology* **37**: 911–917.

Boedigheimer MJ, Wolfinger RD, Bass MB, Bushel PR, Chou JW, Cooper M, Corton JC, Fostel J, Hester S, Lee JS, Liu F, Liu J, Qian H-R, Quackenbush J, Pettit S, Thompson KL. 2008. Sources of variation in baseline gene expression levels from toxicogenomics study control animals across multiple laboratories. *BMC Genomics* **9**: 285.

Brennan LJ, Haukedal J a, Earle JC, Keddie B, Harris HL. 2012. Disruption of redox homeostasis leads to oxidative DNA damage in spermatocytes of Wolbachia-infected Drosophila simulans. *Insect Molecular Biology*.

Charmantier A, Garant D. 2005. Environmental quality and evolutionary potential: lessons from wild populations. *Proceedings. Biological Sciences / The Royal Society* **272**: 1415–1425.

Claypool SM, Koehler CM. 2012. The complexity of cardiolipin in health and disease. *Trends in Biochemical Sciences* **37**: 32–41.

Cline MS, Smoot M, Cerami E, Kuchinsky A, Landys N, Workman C, Christmas R, Avila-Campilo I, Creech M, Gross B, Hanspers K, Isserlin R, Kelley R, Killcoyne S, Lotia S, Maere S, Morris J, Ono K, Pavlovic V, Pico AR, Vailaya A, Wang P-L, Adler A, Conklin BR, Hood L, Kuiper M, Sander C, Schmulevich I, Schwikowski B, Warner GJ, Ideker T, Bader GD. 2007. Integration of biological networks and gene expression data using Cytoscape. *Nature Protocols* **2**: 2366–2382.

Coleman R. 2004. Enzymes of triacylglycerol synthesis and their regulation. *Progress in Lipid Research* **43**: 134–176.

Consortium D12 G. 2007. Evolution of genes and genomes on the Drosophila phylogeny. *Nature* **450**: 203–218.

Cooper B, Hammad L, Fisher N, Karty JA, Montooth KL. 2011. In a variable thermal environment selection favors greater plasticity of cell membranes in Drosophila melanogaster. *Evolution* **66**: 1976–1984.

Cossins a R, Prosser CL. 1978. Evolutionary adaptation of membranes to temperature. *Proceedings of the National Academy of Sciences of the United States of America* **75**: 2040–2043.

Cossins AR, Murray PA, Gracey AY, Logue J, Polley S, Caddick M, Brooks S, Postle T, Maclean N. 2002. The role of desaturases in cold-induced lipid restructuring. *Biochemical Society Transactions* **30**: 1082 – 6.

Cossins AR, Raynard RS. 1987. Adaptive responses of animal cell membranes to temperature. *Symposia of the Society for Experimental Biology* **41**: 95–111.

Cully M, You H, Levine AJ, Mak TW. 2006. Beyond PTEN mutations: the PI3K pathway as an integrator of multiple inputs during tumorigenesis. Nature Reviews. *Cancer* **6**: 184–192.

Devaiah SP, Roth MR, Baughman E, Li M, Tamura P, Jeannotte R, Welti R, Wang X. 2006. Quantitative profiling of polar glycerolipid species from organs of wild-type Arabidopsis and a phospholipase Dalpha1 knockout mutant. *Phytochemistry* **67**: 1907–1924.

Duman JG. 2001. Antifreeze and ice nucleator proteins in terrestrial arthropods. *Annual Review of Physiology* **63**: 327–357.

Fast PG. 1966. A comparative study of the phospholipids and fatty acids of some insects. *Lipids* **1**: 209–215.

Graveley BR, Brooks AN, Carlson JW, Duff MO, Landolin JM, Yang L, Artieri CG, van Baren MJ, Boley N, Booth BW, Brown JB, Cherbas L, Davis CA, Dobin A, Li R, Lin W, Malone JH, Mattiuzzo NR, Miller D, Sturgill D, Tuch BB, Zaleski C, Zhang D, Blanchette M, Dudoit S, Eads B, Green RE, Hammonds A, Jiang L, Kapranov P, Langton L, Perrimon N, Sandler JE, Wan KH, Willingham A, Zhang Y, Zou Y, Andrews J, Bickel PJ, Brenner SE, Brent MR, Cherbas P, Gingeras TR, Hoskins RA, Kaufman TC, Oliver B, Celniker SE. 2011. The developmental transcriptome of Drosophila melanogaster. *Nature* **471**: 473–479.

Greenberg AJ, Hackett SR, Harshman LG, Clark AG. 2010. A hierarchical Bayesian model for a novel sparse partial diallel crossing design. *Genetics* **185**: 361–373.

Gustavsson L. 1995. Phosphatidylethanol formation: specific effects of ethanol mediated via phospholipase D. *Alcohol and Alcoholism* **30**:391-406.

Hammad L a, Cooper BS, Fisher NP, Montooth KL, Karty J a. 2011. Profiling and quantification of Drosophila melanogaster lipids using liquid chromatography/mass spectrometry. *Rapid Communications in Mass Spectrometry* **25**: 2959–2968.

Hazel J, Landrey S. 1988. Time course of thermal adaptation in plasma membranes of trout kidney. II. Molecular species composition. *American Journal of Physiology* 628–634.

Hazel JR. 1995. Thermal adaptation in biological membranes: is homeoviscous adaptation the explanation? *Annual Review of Physiology* **57**: 19–42.

Hazel JR, Williams EE. 1990. The role of alterations in membrane lipid composition in enabling physiological adaptation of organisms to their physical environment. *Progress in Lipid Research* **29**: 167–227.

Hoek JB, Rubin E. 1990. Alcohol and membrane-associated signal transduction. *Alcohol and Alcoholism* **25**: 143–156.

Jombart T, Devillard S, Balloux F. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics* **11**: 94.

Jones HE, Harwood JL, Bowen ID, Griffiths G. 1992. Lipid composition of subcellular membranes from larvae and prepupae of Drosophila melanogaster. *Lipids* **27**: 984–987.

Kharchenko PV, Alekseyenko AA, Schwartz YB, Minoda A, Riddle NC, Ernst J, Sabo PJ, Larschan E, Gorchakov AA, Gu T, Linder-Basso D, Plachetka A, Shanower G, Tolstorukov MY, Luquette LJ, Xi R, Jung YL, Park RW, Bishop EP, Canfield TK, Sandstrom R, Thurman RE, MacAlpine DM, Stamatoyannopoulos JA, Kellis M, Elgin SCR, Kuroda MI, Pirrotta V, Karpen GH, Park PJ. 2011. Comprehensive analysis of the chromatin landscape in Drosophila melanogaster. *Nature* **471**: 480–485.

Koštál V. 2006. Eco-physiological phases of insect diapause. *Journal of Insect Physiology* **52**: 113–127.

Koštál V, Korbelová J, Rozsypal J, Zahradníčková H, Cimlová J, Tomčala A, Šimek P. 2011. Long-term cold acclimation extends survival time at 0°C and modifies the metabolomic profiles of the larvae of the fruit fly Drosophila melanogaster. *PloS One* **6**: e25025.

Koštál V, Tollarová-Borovanská M. 2009. The 70 kDa heat shock protein assists during the repair of chilling injury in the insect, Pyrrhocoris apterus. *PloS One* **4**: e4546.

McElhaney R, Souza K. 1976. The relationship between environmental temperature, cell growth and the fluidity and physical state of the membrane lipids in Bacillus stearothermophilus. *Biochimica Et Biophysica Acta* **443**: 348–359.

Ohtsu T, Kimura MT, Katagiri C. 1998. How Drosophila species acquire cold tolerance--qualitative changes of phospholipids. *European Journal of Biochemistry* **252**: 608–611.

Okano G, Akino T. 1979. Variations in the molecular species of lung phosphatidylglycerol. *Lipids* **14**: 541–546.

Overgaard J, Sørensen JG, Petersen SO, Loeschcke V, Holmstrup M. 2005. Changes in membrane lipid composition following rapid cold hardening in Drosophila melanogaster. *Journal of Insect Physiology* **51**: 1173–1182.

Overgaard J, Tomcala A, Sørensen JG, Holmstrup M, Krogh PH, Simek P, Kostál V. 2008. Effects of acclimation temperature on thermal tolerance and membrane phospholipid composition in the fruit fly Drosophila melanogaster. *Journal of Insect Physiology* **54**: 619–629.

Parisi M, Li R, Oliver B. 2011. Lipid profiles of female and male Drosophila. *BMC Research Notes* **4**: 198.

Pruitt N. 1988. Membrane lipid composition and overwintering strategy in thermally acclimated crayfish. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology* **254***:* 870–876.

Punj V, Chakrabarty a M. 2003. Redox proteins in mammalian cell death: an evolutionarily conserved function in mitochondria and prokaryotes. *Cellular Microbiology* **5**: 225–231.

Rinehart JP, Li A, Yocum GD, Robich RM, Hayward SAL, Denlinger DL. 2007. Up-regulation of heat shock proteins is essential for cold survival during insect diapause. *Proceedings of the National Academy of Sciences of the United States of America* **104**: 11130–11137.

Rubin E, Rottenberg H. 1982. Ethanol-induced injury and adaptation in biological membranes. *Federation Proceedings* **41**: 2465–2471.

Scherer A. 2009. Batch Effects and Noise in Microarray Experiments: Sources and Solutions (Chichester, UK: John Wiley & Sons, Ltd).

Shukla SD, Sun GY, Gibson Wood W, Savolainen MJ, Alling C, Hoek JB. 2001. Ethanol and lipid metabolic signaling. *Alcoholism, Clinical and Experimental Research* **25**: 33S–39S.

Sinclair BJ, Roberts SP. 2005. Acclimation, shock and hardening in the cold. *Journal of Thermal Biology* **30**: 557–562.

Sinensky M. 1974. Homeoviscous adaptation--a homeostatic process that regulates the viscosity of membrane lipids in Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America* **71**: 522–525.

Skwarek LC, Boulianne GL. 2009. Great expectations for PIP: phosphoinositides as regulators of signaling during development and disease. *Developmental Cell* **16**: 12–20.

Sowell RA, Hersberger KE, Kaufman TC, Clemmer DE. 2007. Examining the proteome of Drosophila across organism lifespan. *Journal of Proteome Research* **6**: 3637–3647.

Stone E a, Ayroles JF. 2009. Modulated modularity clustering as an exploratory tool for functional genomic inference. *PLoS Genetics* **5**: e1000479.

Sun G, Sun A. 1985. Ethanol and Membrane Lipids. *Alcoholism: Clinical and Experimental Research* **9**: 164–180.

Taraschi TF, Rubin E. 1985. Effects of ethanol on the chemical and structural properties of biologic membranes. *Laboratory Investigation; a Journal of Technical Methods and Pathology* **52**: 120–131.

Tiku PE, Gracey AY, Macartney AI, Beynon RJ, Cossins AR. 1996. Cold-induced expression of delta 9-desaturase in carp by transcriptional and posttranslational mechanisms. *Science (New York, N.Y.)* **271**: 815–818.

Tomcala a, Tollarová M, Overgaard J, Simek P, Kostál V. 2006. Seasonal acquisition of chill tolerance and restructuring of membrane glycerophospholipids in an overwintering insect: triggering by low temperature, desiccation and diapause progression. *The Journal of Experimental Biology* **209**: 4102–4114.

Vance JE. 2003. Molecular and cell biology of phosphatidylserine and phosphatidylethanolamine metabolism. *Progress in Nucleic Acid Research and Molecular Biology* **75**: 69–111.

Welti R, Li W, Li M, Sang Y, Biesiada H, Zhou H-E, Rajashekar CB, Williams TD, Wang X. 2002. Profiling membrane lipids in plant stress responses. Role of

phospholipase D alpha in freezing-induced lipid changes in Arabidopsis. *The Journal of Biological Chemistry* **277**: 31994–32002.

Zachariassen KE, Kristiansen E. 2000. Ice nucleation and antinucleation in nature. *Cryobiology* **41**: 257–279.

CHAPTER 4

CONCLUSIONS AND FUTURE PERSPECTIVES

4.1 Runx1 in epithelial stem cells

Runx1 has become infamous through its role in hematopoiesis and leukemia. Prior to my study the HFSCs were the only epithelial SC population with a known Runx1 contribution. Runx1 loss in HFSC prevents exit from the quiescent phase resulting in a block of hair cycle progression. Here I presented evidence that labeled Runx1 HFSCs give rise to all HF lineages over extended periods of time, thus these Runx1 expressing HFSCs are truly stem cells. Moreover, I showed that two other epithelial SC populations likely also contain Runx1 expressing cells, intestinal and oral tissue SCs.

Through immunofluorescence staining I determined that a subset of intestinal SCs express Runx1. Lineage tracing these Runx1 positive cells is the crucial next step to identify whether these cells are also functional SCs and can contribute to tissue maintenance for extended periods of time. However, the Runx1CreER mice I used to trace the HF and tumor fate of Runx1 expressing HFSCs in Chapter 2 are not applicable in the intestine, because of its low labeling efficiency when TM is fed to the mice or injected. Thus most likely a new transgenic mouse is necessary to analyze if Runx1 expressing cells in the intestinal crypts are truly SCs.

Additionally, I bring functional evidence that Runx1 marks oral tissue SCs. Runx1 is expressed in some basal layer cells of oral epithelium and these cells contribute

to all layers over extended periods of time. Thus there now are three tissues where there is functional evidence that Runx1 is a SC factor (blood, HF, oral tissue), and in a fourth tissue (intestinal epithelium) for which we have evidence that Runx1 is at least expressed in SCs.

The HF niche is a unique environment in the skin and it has been shown that Runx1 loss can be tolerated and the resulting delay in hair cycle progression can be overcome spontaneously. On the other hand Runx1 loss at different stages in the hematopoietic system has more dramatic phenotypes. It will be interesting to assess which response oral and intestinal SCs elicit upon Runx1 loss. To achieve tissue-, or SC-specific ablation of Runx1 expression, one can use Ah-Cre (intestinal tissue) (Sansom et al., 2004), Lgr5-GFP-IRES-CreER (intestinal SCs), K14-CreER (oral tissue); Runx1CreER (oral SCs) mice in combination with $Runx1^{fl/fl}$ mice and evaluate SC behavior and tissue homeostasis in the absence of Runx1. Additionally, this system can be exploited to start understanding the regulatory networks surrounding Runx1 in these SCs. In the HFSCs p63, p21 and Lef1 have been implicated to directly interact with Runx1 and determine HF biology. Using iKO mouse models in the oral and intestinal tissue will help us understand if these interactions of Runx1 are a common axis of SC regulation. Here, I also presented that Runx1 regulates Stat3 activation by transcriptionally repressing Stat3 suppressors in cancer cells and in primary keratinocytes. As I summarized in Chapter 1 Stat3 KO in the skin has a very similar phenotype to Runx1 KO. It is likely that Runx1 is also upstream of Stat3 activation in normal

HFSC biology, but it still has to be confirmed in freshly isolated HFSCs. Interestingly, Stat3 also has Runx1 binding sites and the question remains if they are used in the skin SCs, or if it is same mechanism as the keratinocytes and cancer.

In general terms my findings suggest that Runx1 is a more global SC factor than previously appreciated. To support this hypothesis additional SC niches should be analyzed for Runx1 expression and its functional significance. Interesting targets would be mammary, prostate or ovarian epithelial SC niches because Runx1 has been shown to be involved in cancers of all of these tissues (see Chapter 1) and at least in the HF there is a direct connection between Runx1 expression in the SC niche and cancer development.

4.2 Runx1 in epithelial cancers

Prior to the study I presented in Chapter 2 little was known about Runx1's role in solid tumor formation in any organism. The Tumbar laboratory had previously shown that Runx1 is crucial for mouse skin SCC formation, but did neither know the mechanism nor the timing with which it supported tumor formation. Moreover, it was unclear if findings would translate from the model organism mouse to human cancer. Here I show that the Runx1 expression pattern in human skin and oral SCC is reminiscent of our observations in mouse SCC. Importantly, human and mouse SCC cell lines both fail to survive Runx1 loss. Overall, my data suggests that the Runx1-dependent mechanisms of SCC formation are conserved

between mouse and human, setting the stage to analyze tumor formation in the model organism mouse.

In the mouse I show that Runx1 acts in the HFSCs, but not in a differentiated population of Runx1 expressing HF cells, to allow tumor initiation. In established skin tumors Runx1 is needed to maintain the SC population as marked by CD34 as well as high levels of proliferation as assayed by the absence of activated Stat3 in Runx1 KO tumors. This raises two main questions, 1. Which molecular changes is Runx1 required for in tumor initiating cells? 2. Why are only SCs with Runx1 able to form tumors? Answers to these questions would not only benefit the skin biology field, but also have the potential to be translated to other SC niches that are the origin of cancer, such as the intestine. The lineage tracing system I employed in Chapter 2 can be used to identify the molecular changes important in Runx1-dependent tumor formation. Briefly, Runx1CreER; tdTomato mice are TM injected and labeled cells are extracted five days later, one day after DMBA or acetone treatment. Next, the $Runx1^{fl}$ allele is crossed into this model, resulting in Runx1 KO and cell labeling after TM injection. Similarly, cells are extracted 5 days after DMBA and acetone treatment. Comparing these four populations using microarrays or RNA-sequencing will help to identify the global changes that happen after the initiating mutation has been established and the influence Runx1 presence or absence has on them. Additionally, to contrast SCs versus differentiated cells in the tumor initiating process the Runx1 infundibular population can be added as samples five (DMBA) and six (acetone). One response

that would be expected in SCs lacking Runx1 based on the findings I presented here is the up-regulation of SOCS genes.

Mechanistically, I demonstrated *in vitro* that Runx1's cancer cell survival functions are mediated through suppression of SOCS3 and SOCS4. SOCS stands for "suppressors of cytokine signaling" and these proteins are part of the negative feedback-loop initiated by Stat3 activation in which they are responsible for preventing Stat3 over-activation. Thus, by blocking SOCS expression Runx1 allows for over-activation of Stat3. My *in vitro* data indicate that this mechanism is also in place in oral SCC and ovarian cancer, at least partially. Confirming the role of Runx1 in these malignancies will help us understand the universal potency of Runx1 as a cancer regulator. Combining the LSL-Kras$^{G12D/+}$ mouse with Runx1CreER; tdTomato mice would allow us to asses if the SC population we identified in the oral epithelium can be at the origin of the oral tumors we observed in the Kras mice. In the ovarian system our *in vitro* results first need to be translated to *in vivo*. To this end the Runx1 WT or iKO mice can be combined with p53$^{fl/fl}$; Rb1$^{fl/fl}$ mice (Flesken-Nikitin et al., 2003) to analyze if they form epithelial ovarian cancer. My findings predict that Runx1 KO mice form significantly fewer tumors than wild-type mice. Additionally, a model of ovarian endometroid carcinoma of LSL-Kras^{G12D}; PTEN$^{fl/fl}$ (Dinulescu et al., 2005) could be used to determine if there are differences across tumor subtypes, like it has been observed in breast cancer (see Chapter1). In addition, the expression studies reviewed in Chapter 1 suggest that Runx1 has an important role in

endometroid carcinoma and that it's loss might prevent or slow down tumor growth, two aspects that should be assayed using Runx1 iKO mice.

Lastly, to understand if Runx1 is a truly global cancer regulator one should analyze the additional cancer types that showed Runx1 over-expression compared to their normal tissue in my Oncomine meta-analysis (Chapter 2). Cell lines from all tumor types can be transformed with Runx1 shRNA and cell survival in response to Runx1 loss is assessed after puromycin selection, equivalent to my work in skin and oral SCC cell lines. Cell lines that show survival defects in the absence of Runx1 can be treated with human OSM to determine if the growth defect was mediated through the loss of activated Stat3. Similar to the notion of Runx1 being a global SC regulator, Runx1 might be a global cancer regulator and extending the panel of tumor types analyzed would help us assess its potential.

Although I focused on analyzing Runx1's tumor promoting role in solid cancers, I would like to end here by emphasizing again that Runx1 is not a one trick pony. We have seen different effects of Runx1 expression on tumor formation depending on the tissue type. Additionally, the protein networks that define the mechanisms of Runx1's role seem to be unique to each cancer (sub-) type. Only detailed future studies in each tissue can help us to understand this transcription factor.

Concluding, my research has significantly extended the field of Runx1 function in normal homeostasis and cancer. Now is the time to evaluate if Runx1 is a truly global SC and cancer regulator.

4.3 Lipid profiling in *D. melanogaster*

In the third Chapter of this dissertation I presented exploratory work on lipidomics in *D. melanogaster*. This is the first time that the natural variation in lipid content across ectotherms has been analyzed. Using 5 geographically distinct populations of *D. melanogaster* I showed that lipids are highly heritable. Each population is characterized by a specific lipid profile that also differs across the sexes. Overall, females and males displayed striking differences in concentration differences of selected polar and neutral lipid species, but also in the self-correlation matrix of all lipids. I also detected particular physiological differences in some populations. The lipid profile of Netherlands' flies indicates rigid membranes in these flies and this is confirmed by determining the ethanol tolerance, which is increased with membrane rigidity.

Additionally, I correlated *Wolbachia* abundance with lipid content and determined that odd-chain lipids and PSs are robustly correlated to *Wolbachia*. This suggests the hypothesis that *Wolbachia* is a provider of odd-chained lipids, while potentially increasing the level of apoptosis. Since all PS lipid species were positively correlated with *Wolbachia* abundance, this would also suggest that all PS lipids can label apoptotic cells for clearance and that double bond confirmation and FA length only modulates the signal strength, if at all. This is

first study that investigated the connection between endosymbionts and lipid content and already with this general survey approach we find these strong connections, warranting further research on this topic and on the dataset as a whole.

Like any datasets that extend our knowledge and ventures into a new field the analysis I presented in Chapter 3 lead to more questions that it could answer, especially regarding the biological significance of lipid variation. Do different FA lengths change the signaling behavior of phospholipids, like PS and PI? Is there a biological significance of double bond conformation in these lipids? And lastly, can lipids be used to is uniquely identify a sex or population of a fly and characterize its stress levels?

To analyze the effect of PS variation a subset of flies could be exposed to various stresses and PS species concentration is measured at different time points of recovery. In addition, total apoptosis can be measured by extracting protein from whole-flies and assaying for cleaved capase3. This experiment could provide new insights into the behavior of specific PSs under different conditions, its effect-size and duration as well as its specificity. Moreover, it could expand our general understanding of apoptosis signaling and the involvement of signaling lipids. Additionally, the positive correlation of *Wolbachia* abundance with PSs raises the question if PS levels drop when lines are cleared of *Wolbachia* using tetracycline. These experiments would be straight forward as they could use existing protocols in the laboratory and the lipid extraction method used here and

additionally this dataset can serve as the baseline. This would determine if PS levels are directly decreased by *Wolbachia* depletion or if the initial correlation detected here stems from secondary effects.

To connect the variations in lipid concentrations to variation in lipid processing enzymes we can carry out a SNP association study. As I briefly touched upon in the main text, the Clark lab has already determined genome-wide SNPs for all but a few fly lines used in this study. With this approach one could not only assess population differences, but global line variation. The power of this dataset lies in the sampling of fly lines. Lines are wild-caught and inbred for over 13 generations, resulting in a panel reflecting the natural variation. Combining these two layers of information, lipid concentrations and genomic SNPs, we can aim to sketch and understand the vast network of protein-lipid interactions.

Concluding, this analysis of lipid biology opened up a new field of biological interactions that is now waiting to be explored.

REFERENCES

Flesken-Nikitin A, Choi K-C, Eng JP, Shmidt EN, Nikitin AY. 2003. Induction of carcinogenesis by concurrent inactivation of p53 and Rb1 in the mouse ovarian surface epithelium. Cancer Research 63: 3459–3463.

Dinulescu DM, Ince TA, Quade BJ, Shafer SA, Crowley D, Jacks T. 2005. Role of K-ras and Pten in the development of mouse models of endometriosis and endometrioid ovarian cancer. Nature Medicine 11: 63–70.

Sansom OJ, Reed KR, Hayes AJ, Ireland H, Brinkmann H, Newton IP, Batlle E, Simon-Assmann P, Clevers H, Nathke IS, Clarke AR, Winton DJ. 2004. Loss of Apc in vivo immediately perturbs Wnt signaling, differentiation, and migration. Genes & Development 18: 1385–1390.

Appendix 1: Fold change results for SA Biosciences qPCR Array Plate

Well	Gene Symbol	p-value	Fold up - or down - regulation WT/KO	Status of raw C_T value
F09	Socs4	0.001	2.39	OKAY
A07	Crk	0.002	2.84	OKAY
F08	Socs3	0.005	2.19	OKAY
E08	Saa3	0.025	3.70	OKAY
E02	Pias1	0.011	1.89	OKAY
D02	Jak2	0.015	1.80	OKAY
C11	Irf1	0.029	1.71	OKAY
B03	F2r	0.031	1.57	OKAY
D11	Oas1a	0.047	2.04	OKAY
G07	Stat5b	0.053	1.47	OKAY
E03	Prl	0.057	6.71	B
B04	Fas	0.076	1.64	OKAY
A04	Ccnd1	0.079	1.82	OKAY
A12	Egfr	0.096	1.46	OKAY
E10	Sh2b1	0.107	2.78	OKAY
B02	F2	0.114	-2.24	B
E06	Ptprc	0.116	2.10	B
F11	Sp1	0.116	1.86	OKAY
D01	Jak1	0.126	1.43	OKAY
F01	Smad1	0.136	1.20	OKAY
G12	Yy1	0.138	1.46	OKAY
A09	Csf1r	0.145	1.97	OKAY
G01	Stam	0.146	1.29	OKAY
F06	Socs1	0.170	1.48	OKAY
F03	Smad3	0.171	1.56	OKAY
C04	Il10rb	0.184	1.52	OKAY
B01	Epor	0.187	2.14	B
B10	Ghr	0.195	1.75	OKAY
F07	Socs2	0.204	1.53	OKAY
D10	Nr3c1	0.272	1.60	OKAY
B12	Ifnar1	0.277	1.27	OKAY
A03	Bcl2l1	0.288	1.38	OKAY
A10	Csf2rb2	0.293	2.63	B
E01	Pdgfra	0.304	1.89	OKAY
G04	Stat3	0.306	-1.42	OKAY
B09	Gbp1	0.310	1.92	B
E05	Ptpn1	0.310	-1.50	OKAY

Well	Gene Symbol	p-value	Fold up - or down - regulation WT/KO	Status of raw C_T value
C01	Ifng	0.313	1.51	B
B07	Isg15	0.317	1.49	OKAY
D05	Mmp3	0.320	1.53	B
D07	Myc	0.326	1.41	OKAY
C06	Il2ra	0.340	2.33	B
F05	Smad5	0.342	1.47	OKAY
F02	Smad2	0.350	1.52	OKAY
E07	Pzp	0.393	2.84	B
E04	Prlr	0.394	1.75	B
B06	Fcgr1	0.399	1.28	B
D08	Nfkb1	0.414	1.32	OKAY
C10	Il6st	0.414	-1.24	OKAY
F12	Src	0.433	-1.11	OKAY
E11	Sit1	0.449	1.37	B
G09	Stub1	0.450	1.19	OKAY
A01	A2m	0.460	-1.26	B
A02	Sh2b2	0.467	1.23	OKAY
E09	Sfpi1	0.469	-1.35	OKAY
G11	Usf1	0.478	1.21	OKAY
C08	Il4	0.518	-1.20	B
A05	Cdkn1a	0.522	-1.34	OKAY
A08	Crp	0.528	1.21	C
B05	Fcer1a	0.528	1.21	C
C03	Il10ra	0.528	1.21	C
D12	Osm	0.528	1.21	C
G05	Stat4	0.528	1.21	C
C05	Il20	0.534	1.75	B
C09	Il4ra	0.538	-1.56	OKAY
D04	Junb	0.551	-1.19	OKAY
C02	Ifngr1	0.569	-1.12	OKAY
G06	Stat5a	0.600	1.07	OKAY
E12	Sla2	0.633	1.18	OKAY
D03	Jun	0.638	1.18	OKAY
G02	Stat1	0.675	-1.15	OKAY
C07	Il2rg	0.697	-1.18	OKAY
G08	Stat6	0.715	1.04	OKAY
B08	Gata3	0.718	-1.05	OKAY
F10	Socs5	0.729	1.06	OKAY
B11	Hmga1	0.748	-1.14	OKAY
G10	Tyk2	0.814	1.10	OKAY
A06	Cebpb	0.841	1.06	OKAY

Well	Gene Symbol	p-value	Fold up - or down - regulation WT/KO	Status of raw C_T value
C12	Irf9	0.854	-1.03	OKAY
D09	Nos2	0.856	-1.57	OKAY
F04	Smad4	0.880	-1.04	OKAY
D06	Mpl	0.931	1.30	B
A11	Cxcl9	0.954	1.06	B
G03	Stat2	N/A	-1.35	OKAY

Appendix 2: Raw CT values for SA Biosciences qPCR Array Plate

Well	Gene Symbol	Gene Description	iKO + Ethanol			iKO + TM (4d)		
			1	2	3	1	2	3
A01	A2m	Alpha-2-macroglobulin	33.19	N/A	N/A	36.3	N/A	N/A
A02	Sh2b2	SH2B adaptor protein 2	28.6	28.47	28.17	27.64	27.78	29.73
A03	Bcl2l1	Bcl2-like 1	20	20.96	20.47	19.84	20.81	20.19
A04	Ccnd1	Cyclin D1	20.45	21.13	20.39	19.57	20.18	20.44
A05	Cdkn1a	Cyclin-dependent kinase inhibitor 1A (P21)	18.06	19.58	18.74	18.73	19.66	20.07
A06	Cebpb	CCAAT/enhancer binding protein (C/EBP), beta	19.7	19.73	19.85	19.63	20.59	19.6
A07	Crk	V-crk sarcoma virus CT10 oncogene homolog (avian)	21.44	21.65	21.49	20.29	20.37	20.21
A08	Crp	C-reactive protein, pentraxin-related	N/A	N/A	N/A	N/A	N/A	N/A
A09	Csf1r	Colony stimulating factor 1 receptor	28.57	28.57	28.47	27.19	27.65	28.64
A10	Csf2rb2	Colony stimulating factor 2 receptor, beta 2, low-affinity (granulocyte-macrophage)	N/A	N/A	N/A	N/A	32.15	34.47
A11	Cxcl9	Chemokine (C-X-C motif) ligand 9	34.45	N/A	N/A	N/A	37.87	N/A
A12	Egfr	Epidermal growth factor receptor	22.8	23.55	22.83	22.45	22.95	22.96
B01	Epor	Erythropoietin receptor	30.05	32.33	31.5	30.12	30.23	31.05
B02	F2	Coagulation factor II	33.27	32.78	32.57	34.05	36.92	33.88
B03	F2r	Coagulation factor II (thrombin) receptor	23.63	23.93	23.54	23.43	23.33	23.21
B04	Fas	Fas (TNF receptor superfamily member 6)	26.47	26.7	26.8	26.34	25.95	26.35
B05	Fcer1a	Fc receptor, IgE, high affinity I, alpha polypeptide	N/A	N/A	N/A	36.62	N/A	N/A
B06	Fcgr1	Fc receptor, IgG, high affinity I	N/A	N/A	35.25	35.46	37.87	34.75
B07	Isg15	ISG15 ubiquitin-like modifier	27.08	27.57	27.96	27.73	27.09	26.89
B08	Gata3	GATA binding protein 3	23.49	23.7	23.89	23.81	24.2	24.11
B09	Gbp1	Guanylate binding protein 1	32.95	36.79	31.7	32.06	34.01	31.57
B10	Ghr	Growth hormone receptor	25.09	25.07	25.19	23.97	24.7	25.06
B11	Hmga1	High mobility group AT-hook 1	19.8	20.39	20.21	20.32	20.02	21.43

Well	Gene Symbol	Gene Description	iKO + Ethanol			iKO + TM (4d)		
B12	Ifnar1	Interferon (alpha and beta) receptor 1	21.83	22.08	21.81	21.61	21.92	21.98
C01	Ifng	Interferon gamma	N/A	N/A	N/A	N/A	N/A	34.04
C02	Ifngr1	Interferon gamma receptor 1	21.46	21.4	20.82	21.39	21.59	21.99
C03	Il10ra	Interleukin 10 receptor, alpha	N/A	N/A	N/A	N/A	N/A	N/A
C04	Il10rb	Interleukin 10 receptor, beta	22.75	22.61	21.02	21.74	21.85	21.8
C05	Il20	Interleukin 20	N/A	33.17	35.11	31.56	N/A	N/A
C06	Il2ra	Interleukin 2 receptor, alpha chain	N/A	N/A	36.79	N/A	32.15	N/A
C07	Il2rg	Interleukin 2 receptor, gamma chain	26.41	27.38	25.8	26.74	27.73	26.63
C08	Il4	Interleukin 4	30.62	32.6	N/A	N/A	32.16	32.67
C09	Il4ra	Interleukin 4 receptor, alpha	27.67	28.29	25.7	28.41	28.94	27.04
C10	Il6st	Interleukin 6 signal transducer	22.68	23.89	22.57	23.59	24.06	23.24
C11	Irf1	Interferon regulatory factor 1	25.04	25.53	24.69	24.3	24.49	24.96
C12	Irf9	Interferon regulatory factor 9	24.15	24.22	23.24	23.94	24.04	24.58
D01	Jak1	Janus kinase 1	21.39	21.25	20.9	20.9	21.01	20.88
D02	Jak2	Janus kinase 2	23.63	23.94	22.86	22.67	22.86	23.18
D03	Jun	Jun oncogene	21.44	22.08	20.09	21.46	21.1	21.15
D04	Junb	Jun-B oncogene	20.36	21.94	19.68	20.79	21.64	21.11
D05	Mmp3	Matrix metallopeptidase 3	30.45	32.08	29.75	31.44	30.05	29.77
D06	Mpl	Myeloproliferative leukemia virus oncogene	37.65	32.17	N/A	37.94	35.09	31.86
D07	Myc	Myelocytomatosis oncogene	18.8	19.11	17.31	18.18	18.58	17.77
D08	Nfkb1	Nuclear factor of kappa light polypeptide gene enhancer in B-cells 1, p105	22.1	22.42	20.73	21.62	22.16	21.07
D09	Nos2	Nitric oxide synthase 2, inducible	25.06	27.61	26.79	27.77	28.97	25.49
D10	Nr3c1	Nuclear receptor subfamily 3, group C, member 1	22.14	22.55	20.93	22.62	21.09	20.7
D11	Oas1a	2'-5' oligoadenylate synthetase 1A	26.2	26.01	25.78	25.72	24.88	25.12
D12	Osm	Oncostatin M	N/A	N/A	N/A	N/A	N/A	N/A

Well	Gene Symbol	Gene Description	iKO + Ethanol			iKO + TM (4d)		
E01	Pdgfra	Platelet derived growth factor receptor, alpha polypeptide	27.74	29.25	27.22	28.29	27.7	26.27
E02	Pias1	Protein inhibitor of activated STAT 1	22.34	23.51	22.5	22.07	21.9	22.44
E03	Prl	Prolactin	N/A	N/A	N/A	32.28	33.51	31.78
E04	Prlr	Prolactin receptor	31.86	N/A	32.29	30.5	32.05	N/A
E05	Ptpn1	Protein tyrosine phosphatase, non-receptor type 1	19.8	21.82	20.59	21.43	21.86	21.48
E06	Ptprc	Protein tyrosine phosphatase, receptor type, C	36.99	N/A	N/A	N/A	33.92	33.68
E07	Pzp	Pregnancy zone protein	33.77	32.61	29.52	30.05	30.98	31.17
E08	Saa3	Serum amyloid A 3	26.89	26.83	26.96	25.22	25.74	24.87
E09	Sfpi1	SFFV proviral integration 1	27.82	28.62	25.76	28.62	27.88	27.81
E10	Sh2b1	SH2B adaptor protein 1	23.73	23.67	23.87	21.84	22.13	23.68
E11	Sit1	Suppression inducing transmembrane adaptor 1	34.24	35.61	36.42	33.78	34.91	35.29
E12	Sla2	Src-like-adaptor 2	27.5	28.79	28.31	27.87	28.92	27.92
F01	Smad1	MAD homolog 1 (Drosophila)	21.4	21.93	20.59	21.41	21.07	21.45
F02	Smad2	MAD homolog 2 (Drosophila)	21.94	23.34	20.84	21.57	22.44	21.1
F03	Smad3	MAD homolog 3 (Drosophila)	23.14	24.67	22.25	23.06	22.7	23.2
F04	Smad4	MAD homolog 4 (Drosophila)	20.06	20.66	19.8	20.03	20.04	21.42
F05	Smad5	MAD homolog 5 (Drosophila)	21.11	22.72	22.12	21.44	21.48	22.17
F06	Socs1	Suppressor of cytokine signaling 1	27.07	27.74	26.04	26.51	27.08	26.37
F07	Socs2	Suppressor of cytokine signaling 2	23.87	23.95	22.16	23.04	23.07	22.84
F08	Socs3	Suppressor of cytokine signaling 3	25.77	25.78	25.32	24.92	24.71	24.65
F09	Socs4	Suppressor of cytokine signaling 4	24.09	24.95	23.49	23.3	23.16	23.11
F10	Socs5	Suppressor of cytokine signaling 5	22.47	22.72	21.56	21.94	22.67	22.69
F11	Sp1	Trans-acting transcription factor 1	21.62	21.39	19.64	19.99	20.4	20.39
F12	Src	Rous sarcoma oncogene	20.3	21.08	19.75	20.6	20.96	20.82
G01	Stam	Signal transducing adaptor molecule (SH3 domain and ITAM motif) 1	20.99	22.17	20.54	21.3	21.12	20.98
G02	Stat1	Signal transducer and activator of transcription 1	23.38	24.13	23.18	23.51	23.97	24.63

Well	Gene Symbol	Gene Description	iKO + Ethanol			iKO + TM (4d)		
G03	Stat2	Signal transducer and activator of transcription 2	22.32	23.61	22.43	23.66		23.41
G04	Stat3	Signal transducer and activator of transcription 3	20.64	21.95	21.32	21.86	21.63	22.74
G05	Stat4	Signal transducer and activator of transcription 4	35.61	N/A	N/A	N/A	N/A	N/A
G06	Stat5a	Signal transducer and activator of transcription 5A	24.36	25.36	24.04	25.12	24.56	24.6
G07	Stat5b	Signal transducer and activator of transcription 5B	21.83	22.53	21.49	21.99	21.58	21.42
G08	Stat6	Signal transducer and activator of transcription 6	20.28	20.76	20.09	20.72	20.49	20.57
G09	Stub1	STIP1 homology and U-Box containing protein 1	19.59	20.7	18.91	19.67	20.1	19.5
G10	Tyk2	Tyrosine kinase 2	23.95	23.83	22.33	23.7	23.8	23.01
G11	Usf1	Upstream transcription factor 1	23.53	23.27	22.59	22.86	22.77	23.74
G12	Yy1	YY1 transcription factor	19.66	20.33	19.91	19.97	19.73	19.37
H01	Gusb	Glucuronidase, beta	22.99	23.35	22.62	22.99	22.86	23.04
H02	Hprt	Hypoxanthine guanine phosphoribosyl transferase	19.3	19.93	18.99	19.68	18.9	20.37
H03	Hsp90ab1	Heat shock protein 90 alpha (cytosolic), class B member 1	16.05	16.08	15.87	16.08	15.7	15.83
H04	Gapdh	Glyceraldehyde-3-phosphate dehydrogenase	15.84	16.64	14.84	16.44	16.68	15.44
H05	Actb	Actin, beta	14.83	16.31	14.72	15.77	16.18	16.46
H06	MGDC	Mouse Genomic DNA Contamination	N/A	N/A	N/A	N/A	N/A	38.43
H07	RTC	Reverse Transcription Control	N/A	N/A	N/A	N/A	N/A	N/A
H08	RTC	Reverse Transcription Control	N/A	N/A	N/A	N/A	N/A	N/A
H09	RTC	Reverse Transcription Control	N/A	N/A	N/A	N/A	N/A	N/A
H10	PPC	Positive PCR Control	16.38	15.95	15.95	16.14	16.21	16.06
H11	PPC	Positive PCR Control	16.16	16.27	16.08	16.09	16.53	16.68
H12	PPC	Positive PCR Control	16.42	16.37	16.24	16.59	16.41	16.8

Appendix 3: Complete list of lipid species analyzed in Chapter 3

Lipid Class	Lipid Species	Alternate Identification	Mass	Formula
DAG	**DAG16:1_30:1**		556.49237	C33H66O5N
DAG	**DAG18:1_30:1**		556.49237	C33H66O5N
DAG	**DAG16:1_30:2**		554.47677	C33H64O5N
DAG	**DAG18:1_30:2**		554.47677	C33H64O5N
DAG	**DAG16:1_32:1**		584.52357	C35H70O5N
DAG	**DAG18:1_32:1**		584.52357	C35H70O5N
DAG	**DAG16:1_32:2**		582.50797	C35H68O5N
DAG	**DAG18:1_32:2**		582.50797	C35H68O5N
DAG	**DAG16:1_33:1**		598.53917	C36H72O5N
DAG	**DAG18:1_33:1**		598.53917	C36H72O5N
DAG	**DAG16:1_33:2**		596.52357	C36H70O5N
DAG	**DAG16:1_34:1**		612.55477	C37H74O5N
DAG	**DAG18:1_34:1**		612.55477	C37H74O5N
DAG	**DAG16:1_34:2**		610.53917	C37H72O5N
DAG	**DAG18:1_34:2**		610.53917	C37H72O5N
DAG	**DAG16:1_34:3**		608.52357	C37H70O5N
DAG	**DAG16:1_35:1**		626.57037	C38H76O5N
DAG	**DAG18:1_35:1**		626.57037	C38H76O5N
DAG	**DAG16:1_35:2**		624.55477	C38H74O5N
DAG	**DAG16:1_36:1**		640.58597	C39H78O5N
DAG	**DAG18:1_36:1**		640.58597	C39H78O5N
DAG	**DAG16:1_36:2**		638.57037	C39H76O5N
DAG	**DAG18:1_36:2**		638.57037	C39H76O5N
DAG	**DAG18:1_36:3**		636.55477	C39H74O5N
DAG	**DAG16:1_37:1**		654.60157	C40H80O5N
DAG	**DAG18:1_37:1**		654.60157	C40H80O5N
DAG	**DAG16:1_37:2**		652.58597	C40H78O5N
DAG	**DAG18:1_37:2**		652.58597	C40H78O5N
DAG	**DAG16:1_37:3**		650.57037	C40H76O5N
DAG	**DAG16:1_38:1**		668.61717	C41H82O5N
DAG	**DAG18:1_38:1**		668.61717	C41H82O5N
DAG	**DAG16:1_38:2**		666.60157	C41H80O5N
DAG	**DAG18:1_38:2**		666.60157	C41H80O5N
TAG	**TAG16:1_40:1**		710.62767	C43H84O6N
TAG	**TAG18:1_40:1**		710.62767	C43H84O6N
TAG	**TAG16:1_40:2**		708.61207	C43H82O6N
TAG	**TAG18:1_40:2**		708.61207	C43H82O6N

Lipid Class	Lipid Species	Alternate Identification	Mass	Formula
TAG	**TAG16:1_41:1**		724.64327	C44H86O6N
TAG	**TAG16:1_41:2**		722.62767	C44H84O6N
TAG	**TAG16:1_42:1**		738.65887	C45H88O6N
TAG	**TAG18:1_42:1**		738.65887	C45H88O6N
TAG	**TAG16:1_42:2**		736.64327	C45H86O6N
TAG	**TAG18:1_42:2**		736.64327	C45H86O6N
TAG	**TAG16:1_43:1**		752.67447	C46H90O6N
TAG	**TAG18:1_43:1**		752.67447	C46H90O6N
TAG	**TAG16:1_43:2**		750.65887	C46H88O6N
TAG	**TAG16:1_44:1**		766.69007	C47H92O6N
TAG	**TAG18:1_44:1**		766.69007	C47H92O6N
TAG	**TAG16:1_44:2**		764.67447	C47H90O6N
TAG	**TAG18:1_44:2**		764.67447	C47H90O6N
TAG	**TAG16:1_45:1**		780.70567	C48H94O6N
TAG	**TAG18:1_45:1**		780.70567	C48H94O6N
TAG	**TAG16:1_45:2**		778.69007	C48H92O6N
TAG	**TAG18:1_45:2**		778.69007	C48H92O6N
TAG	**TAG16:1_45:3**		776.67447	C48H90O6N
TAG	**TAG16:1_46:1**		794.72127	C49H96O6N
TAG	**TAG18:1_46:1**		794.72127	C49H96O6N
TAG	**TAG16:1_46:2**		792.70567	C49H94O6N
TAG	**TAG18:1_46:2**		792.70567	C49H94O6N
TAG	**TAG16:1_46:3**		790.69007	C49H92O6N
TAG	**TAG18:1_46:3**		790.69007	C49H92O6N
TAG	**TAG16:1_47:1**		808.73687	C50H98O6N
TAG	**TAG18:1_47:1**		808.73687	C50H98O6N
TAG	**TAG16:1_47:2**		806.72127	C50H96O6N
TAG	**TAG18:1_47:2**		806.72127	C50H96O6N
TAG	**TAG16:1_47:3**		804.70567	C50H94O6N
TAG	**TAG16:1_48:1**		822.75247	C51H100O6N
TAG	**TAG18:1_48:1**		822.75247	C51H100O6N
TAG	**TAG16:1_48:2**		820.73687	C51H98O6N
TAG	**TAG18:1_48:2**		820.73687	C51H98O6N
TAG	**TAG16:1_48:3**		818.72127	C51H96O6N
TAG	**TAG18:1_48:3**		818.72127	C51H96O6N
TAG	**TAG16:1_49:1**		836.76807	C52H102O6N
TAG	**TAG18:1_49:1**		836.76807	C52H102O6N
TAG	**TAG16:1_49:2**		834.75247	C52H100O6N

Lipid Class	Lipid Species	Alternate Identification	Mass	Formula
TAG	**TAG18:1_49:2**		834.75247	C52H100O6N
TAG	**TAG16:1_49:3**		832.73687	C52H98O6N
TAG	**TAG18:1_49:3**		832.73687	C52H98O6N
TAG	**TAG16:1_50:1**		850.78367	C53H104O6N
TAG	**TAG18:1_50:1**		850.78367	C53H104O6N
TAG	**TAG16:1_50:2**		848.76807	C53H102O6N
TAG	**TAG18:1_50:2**		848.76807	C53H102O6N
TAG	**TAG16:1_50:3**		846.75247	C53H100O6N
TAG	**TAG18:1_50:3**		846.75247	C53H100O6N
TAG	**TAG16:1_50:4**		844.73687	C53H98O6N
TAG	**TAG16:1_51:1**		864.79927	C54H106O6N
TAG	**TAG18:1_51:1**		864.79927	C54H106O6N
TAG	**TAG16:1_52:2**		876.79927	C55H106O6N
TAG	**TAG18:1_52:2**		876.79927	C55H106O6N
TAG	**TAG16:1_52:3**		874.78367	C55H104O6N
TAG	**TAG18:1_52:3**		874.78367	C55H104O6N
TAG	**TAG18:1_53:2**		890.81487	C56H108O6N
TAG	**TAG18:1_53:3**		888.79927	C56H106O6N
TAG	**TAG18:1_54:3**		902.81487	C57H108O6N
lysoPC	**lysoPC 14:0**		468.30767	C22H47O7PN
lysoPC	**lysoPC 14:1**		466.29207	C22H45O7PN
lysoPC	**lysoPC 15:0**		482.32327	C23H49O7PN
lysoPC	**lysoPC 16:0**		496.33887	C24H51O7PN
lysoPC	**lysoPC 16:1**		494.32327	C24H49O7PN
lysoPC	**lysoPC 17:0**		510.35447	C25H53O7PN
lysoPC	**lysoPC 17:1**		508.33887	C25H51O7PN
lysoPC	**lysoPC 18:1**		522.35447	C26H53O7PN
lysoPC	**lysoPC 18:2**		520.33887	C26H51O7PN
lysoPE	**lysoPE 15:0**		440.27647	C20H43O7PN
lysoPE	**lysoPE 16:0**		454.29207	C21H45O7PN
lysoPE	**lysoPE 16:1**		452.27647	C21H43O7PN
lysoPE	**lysoPE 17:0**		468.30767	C22H47O7PN
lysoPE	**lysoPE 17:1**		466.29207	C22H45O7PN
lysoPE	**lysoPE 18:1**		480.30767	C23H47O7PN
lysoPE	**lysoPE 18:2**		478.29207	C23H45O7PN
lysoPE	**lysoPE 20:0**		510.35447	C25H53O7PN
PA	**PA 34:1**		692.52097	C37H75O8PN
PA	**PA 36:2**		718.53657	C39H77O8PN

Lipid Class	Lipid Species	Alternate Identification	Mass	Formula
PC	**PC 28:0**		678.50537	C36H73O8PN
PC	**PC 28:1**		676.48977	C36H71O8PN
PC	**PC 30:0**		706.53657	C38H77O8PN
PC	**PC 30:1**		704.52097	C38H75O8PN
PC	**PC 30:2**		702.50537	C38H73O8PN
PC	**PC 31:0**		720.55217	C39H79O8PN
PC	**PC 31:1**		718.53657	C39H77O8PN
PC	**PC 31:2**		716.52097	C39H75O8PN
PC	**PC 32:0**		734.56777	C40H81O8PN
PC	**PC 32:1**		732.55217	C40H79O8PN
PC	**PC 32:2**		730.53657	C40H77O8PN
PC	**PC 32:3**		728.52097	C40H75O8PN
PC	**PC 33:0**		748.58337	C41H83O8PN
PC	**PC 33:1**		746.56777	C41H81O8PN
PC	**PC 33:2**		744.55217	C41H79O8PN
PC	**PC 33:3**		742.53657	C41H77O8PN
PC	**PC 34:0**		762.59897	C42H85O8PN
PC	**PC 34:1**		760.58337	C42H83O8PN
PC	**PC 34:2**		758.56777	C42H81O8PN
PC	**PC 34:3**		756.55217	C42H79O8PN
PC	**PC 34:4**		754.53657	C42H77O8PN
PC	**PC 35:1**		774.59897	C43H85O8PN
PC	**PC 35:2**		772.58337	C43H83O8PN
PC	**PC 35:3**		770.56777	C43H81O8PN
PC	**PC 35:4**		768.55217	C43H79O8PN
PC	**PC 36:0**		790.63017	C44H89O8PN
PC	**PC 36:1**		788.61457	C44H87O8PN
PC	**PC 36:2**		786.59897	C44H85O8PN
PC	**PC 36:3**		784.58337	C44H83O8PN
PC	**PC 36:4**		782.56777	C44H81O8PN
PC	**PC 37:1**		802.63017	C45H89O8PN
PC	**PC 37:2**		800.61457	C45H87O8PN
PC	**PC 38:1**		816.64577	C46H91O8PN
PC	**PC 38:2**		814.63017	C46H89O8PN
PE	**PE 28:0**		636.45857	C33H67O8PN
PE	**PE 28:1**		634.44297	C33H65O8PN
PE	**PE 30:0**		664.48977	C35H71O8PN
PE	**PE 30:1**		662.47417	C35H69O8PN

Lipid Class	Lipid Species	Alternate Identification	Mass	Formula
PE	**PE 31:0**		678.50537	C36H73O8PN
PE	**PE 31:1**		676.48977	C36H71O8PN
PE	**PE 31:2**		674.47417	C36H69O8PN
PE	**PE 32:0**		692.52097	C37H75O8PN
PE	**PE 32:1**		690.50537	C37H73O8PN
PE	**PE 32:2**		688.48977	C37H71O8PN
PE	**PE 32:3**		686.47417	C37H69O8PN
PE	**PE 33:1**		704.52097	C38H75O8PN
PE	**PE 33:2**		702.50537	C38H73O8PN
PE	**PE 33:3**		700.48977	C38H71O8PN
PE	**PE 34:0**		720.55217	C39H79O8PN
PE	**PE 34:1**		718.53657	C39H77O8PN
PE	**PE 34:2**		716.52097	C39H75O8PN
PE	**PE 34:3**		714.50537	C39H73O8PN
PE	**PE 34:4**		712.48977	C39H71O8PN
PE	**PE 35:1**		732.55217	C40H79O8PN
PE	**PE 35:2**		730.53657	C40H77O8PN
PE	**PE 35:3**		728.52097	C40H75O8PN
PE	**PE 36:1**		746.56777	C41H81O8PN
PE	**PE 36:2**		744.55217	C41H79O8PN
PE	**PE 36:3**		742.53657	C41H77O8PN
PE	**PE 36:4**		740.52097	C41H75O8PN
PE	**PE 37:1**		760.58337	C42H83O8PN
PE	**PE 37:2**		758.56777	C42H81O8PN
PE	**PE 37:3**		756.55217	C42H79O8PN
PE	**PE 38:1**		774.59897	C43H85O8PN
PE	**PE 38:2**		772.58337	C43H83O8PN
PE	**PE 38:5**		766.53657	C43H77O8PN
PE	**PE-Cer 16:0**		661.52644	C36H74O6PN2
PG	**PG 24:0**	TAG 34:0	628.41717	C30H63O10PN
PG	**PG 26:0**	TAG 36:0	656.44837	C32H67O10PN
PG	**PG 26:1**	TAG 36:1	654.43277	C32H65O10PN
PG	**PG 27:0**	TAG 37:0	670.46397	C33H69O10PN
PG	**PG 28:0**	TAG 38:0	684.47957	C34H71O10PN
PG	**PG 28:1**	TAG 38:1	682.46397	C34H69O10PN
PG	**PG 29:0**	TAG 39:0	698.49517	C35H73O10PN
PG	**PG 30:0**	TAG 40:0	712.51077	C36H75O10PN
PG	**PG 30:1**	TAG 40:1	710.49517	C36H73O10PN

Lipid Class	Lipid Species	Alternate Identification	Mass	Formula
PG	**PG 30:2**	TAG 40:2	708.47957	C36H71O10PN
PG	**PG 31:0**	TAG 41:0	726.52637	C37H77O10PN
PG	**PG 31:1**	TAG 41:1	724.51077	C37H75O10PN
PG	**PG 32:0**	TAG 42:0	740.54197	C38H79O10PN
PG	**PG 32:1**	TAG 42:1	738.52637	C38H77O10PN
PG	**PG 32:2**	TAG 42:2	736.51077	C38H75O10PN
PG	**PG 33:1**	TAG 43:1	752.54197	C39H79O10PN
PG	**PG 33:2**	TAG 43:2	750.52637	C39H77O10PN
PG	**PG 34:1**	TAG 44:1	766.55757	C40H81O10PN
PG	**PG 34:2**	TAG 44:2	764.54197	C40H79O10PN
PG	**PG 34:3**	TAG 44:3	762.52637	C40H77O10PN
PG	**PG 36:2**	TAG 46:2	792.57317	C42H83O10PN
PG	**PG 36:3**	TAG 46:3	790.55757	C42H81O10PN
PI	**PI 30:0**		800.52667	C39H79O13PN
PI	**PI 31:0**		814.54227	C40H81O13PN
PI	**PI 31:1**		812.52667	C40H79O13PN
PI	**PI 32:0**		828.55787	C41H83O13PN
PI	**PI 32:1**		826.54227	C41H81O13PN
PI	**PI 32:2**		824.52667	C41H79O13PN
PI	**PI 33:1**		840.55787	C42H83O13PN
PI	**PI 33:2**		838.54227	C42H81O13PN
PI	**PI 34:1**		854.57347	C43H85O13PN
PI	**PI 34:2**		852.55787	C43H83O13PN
PI	**PI 34:3**		850.54227	C43H81O13PN
PI	**PI 35:1**		868.58907	C44H87O13PN
PI	**PI 35:2**		866.57347	C44H85O13PN
PI	**PI 36:1**		882.60467	C45H89O13PN
PI	**PI 36:2**		880.58907	C45H87O13PN
PI	**PI 36:3**		878.57347	C45H85O13PN
PI	**PI 36:4**		876.55787	C45H83O13PN
PI	**PI 38:2**		908.62027	C47H91O13PN
PS	**PS 32:1**		734.49517	C38H73O10PN
PS	**PS 32:2**		732.47957	C38H71O10PN
PS	**PS 34:1**		762.52637	C40H77O10PN
PS	**PS 34:2**		760.51077	C40H75O10PN
PS	**PS 34:3**		758.49517	C40H73O10PN
PS	**PS 34:4**		756.47957	C40H71O10PN
PS	**PS 36:1**		790.55757	C42H81O10PN

Lipid Class	Lipid Species	Alternate Identification	Mass	Formula
PS	**PS 36:2**		788.54197	C42H79O10PN
PS	**PS 36:3**		786.52637	C42H77O10PN
PS	**PS 36:4**		784.51077	C42H75O10PN
PS	**PS 36:5**		782.49517	C42H73O10PN
PS	**PS 38:1**		818.58877	C44H85O10PN
PS	**PS 38:2**		816.57317	C44H83O10PN